JOURNAL OF ATMOSPHERIC AND TERRESTRIAL PHYSICS

EDITORIAL ADVISORY BOARD

H. O. G. ALFVÉN-STOCKHOLM · **J. BARTELS**-GÖTTINGEN · **P. M. S. BLACKETT**-MANCHESTER · **E. G. BOWEN**-SYDNEY · **SIR DAVID BRUNT**-LONDON · **R. BUREAU**-BAGNEUX · **S. CHAPMAN**-OXFORD · **J. EGEDAL**-COPENHAGEN · **H. FICKER**-VIENNA **V. F. HESS**-NEW YORK · **E. O. HULBURT**-WASHINGTON, D.C. · **H. ISRAEL**-TÜBINGEN **SIR NELSON JOHNSON**-LONDON · **D. H. MENZEL**-CAMBRIDGE, MASS. · **S. K. MITRA**-CALCUTTA · **M. NICOLET**-UCCLE · **E. H. NORINDER**-UPSALA · **B. van der POL**-GENEVA · **J. A. RATCLIFFE**-CAMBRIDGE · **E. REGENER**-WEISSENAU (BODENSEE) **O. E. H. RYDBECK**-GOTHENBURG · **SIR GEORGE SIMPSON**-LONDON · **H. SOLBERG**-OSLO · **C. STØRMER**-OSLO · **H. U. SVERDRUP**-OSLO · **J. VELDKAMP**-DE BILT **M. WALDMEIER**-ZÜRICH · **R. v. d. R. WOOLLEY**-CANBERRA **J. ZENNECK**-ALTHEGNENBERG

VOL.
2
1952

EDITOR IN CHIEF

SIR EDWARD APPLETON, F.R.S.
EDINBURGH

VOLUME 2
1952

PERGAMON PRESS LTD · LONDON

VO
2
19

Contents

No. 1

Page

D. R. Bates and H. S. W. Massey: The negative ion concentration in the lower ionosphere . . . 1
T. H. Ellison: The propagation of sound waves through a medium with very small random variations in refractive index . . . 14
Karl Försterling and Hans-Otto Wüster: Über die Entstehung von Oberwellen in der Ionosphäre . . . 22
H. W. Kasemir: An apparatus for simultaneous registration of potential gradient and air-earth current . . . 32
K. Rawer: Comparaison des résultats de mesures de l'absorption ionosphérique effectuées par deux stations européennes . . . 38
K. Bibl and K. Rawer: Les contributions des régions *D* et *E* dans les mesures de l'absorption ionosphérique . . . 51
B. W. Osborne: Ionospheric behaviour in the F_2 region at Singapore . . . 66
Dietrich Stranz: Weitere Ergebnisse der Fade-out-Statistik . . . 79
Book reviews
A. C. B. Lovell: Cosmical Electrodynamics by H. Alfvén . . . 83
R. L. Smith-Rose: Micro-wave Measurements by H. M. Barlow and A. L. Cullen . . . 83

No. 2

D. H. Shinn and H. A. Whale: Group velocities and group heights from the magneto-ionic theory 85
F. H. Northover: The anomalous propagation of radio waves in the 1–10 metre band . . . 106
Sylvan Rubin: Liquid particles in atmospheric haze . . . 130

No. 3

G. J. Phillips: Measurement of winds in the ionosphere . . . 141
J. Alan Chalmers: Negative electric fields in mist and fog . . . 155
W. Jackson: World-wide simultaneous magnetic fluctuations and their relation to sudden commencements . . . 160
Erich Regener: Über Schwankungen des Ozons in der Troposphäre und Stratosphäre (Fluctuations of ozone in the troposphere and stratosphere) . . . 173
H. K. Paetzold: Erfassung der vertikalen Ozonverteilung in verschiedenen geographischen Breiten bei Mondfinsternissen (Determination of vertical ozone distribution at various latitudes during eclipses of the moon) . . . 183
Alfred Ehmert: Gleichzeitige Messungen des Ozongehaltes bodennaher Luft an mehreren Stationen mit einem einfachen absoluten Verfahren (Simultaneous measurements of the ozone content of air near the earth at several stations by means of a simple absolute method) . . . 189
Indices of Geomagnetic Activity of the Observatories Abinger, Eskdalemuir and Lerwick, August to December 1951 . . . 196
Book reviews . . . 199
Corrections . . . 200

No. 4

G. A. O'Donnell: Electric conductivity and small ion concentration of the atmosphere at 1 metre above ground and conductivity at ground level . . . 201
W. C. Bain and R. N. Bracewell: An explanation of radio propagation at 16 Kc/s in terms of two layers below *E* layer . . . 216
R. N. Bracewell: Theory of formation of an ionospheric layer below *E* layer, based on eclipse and solar flare effects at 16 Kc/s . . . 226
E. V. Appleton and W. R. Piggott: The morphology of storms in the F_2 layer of the ionosphere. I. Some statistical relationships . . . 236
Research Notes:
D. R. Bates and H. S. W. Massey: On negative ions of molecular oxygen in the *D* layer . . . 253
B. Landmark: Polarization of radio waves reflected from the ionosphere . . . 254
Indices of Geomagnetic Activity of the Observatories Abinger, Eskdalemuir and Lerwick, January to March 1952 . . . 256
Obituary: Dr. Bernard Lyot . . . 259
Book review:
R. L. Smith-Rose: Recent advances in radio receivers by L. A. Moxon . . . 260

VOL. 2 1952

No. 5

C. M. MINNIS: The graphical representation of the longitude effect in F_2-region 261
P. J. NOLAN and P. J. KENNY: A modified McCLELLAND method for measuring ionic mobilities . 266
K. G. BUDDEN and G. G. YATES: A search for radio echoes of long delay 272
J. S. GREENHOW: A radio echo method for the investigation of atmospheric winds at altitudes of 80 to 100 km 282
J. A. CHALMERS: The relation between point discharge current and field 292
J. A. CHALMERS: Point discharge currents 301
Indices of Geomagnetic Activity of the Observatories ABINGER, ESKDALEMUIR and LERWICK, April to June 1952. 306
Book review:
Latin Treatises on Comets between 1238 and 1368 A.D.; edited by LYNN THORNDIKE
Reviewed by A. C. B. LOVELL 308

No. 6

S. DEB: Nitrogen in the upper atmosphere 309
N. DAVIDS: Theoretical group heights of reflection 150 kc/s radio waves vertically incident on the ionosphere 324
J. A. CHALMERS: Electric charges from ice friction 337
W. DIEMINGER: Über die Ursache der excessiven Absorption in der Ionosphäre an Wintertagen . . 340
F. G. SMITH: Ionospheric refraction of 81·5 Mc/sec radio waves from radio stars 350
C. G. LITTLE and A. MAXWELL: Scintillation of radio stars during aurorae and magnetic storms 356
Research Notes:
J. D. WHITEHEAD: The quasi-transverse (Q.T.) approximation to APPLETONS magneto-ionic equation 361
F. LIED: A physical interpretation of the corrected quasi-transverse approximation for the ordinary wave 362
Indices of Geomagnetic Activity of the Observatories ABINGER, ESKDALEMUIR and LERWICK, July and August 1952 363

Author Index to Volume 2

(B) denotes a book review

APPLETON, E. V. pp. 199 (B), 236.
BAIN, W. C. p. 216.
BATES, D. R. pp. 1, 253.
BIBL, K. p. 51.
BRACEWELL, R. N. pp. 216, 226.
BUDDEN, K. G. p. 272.
CHALMERS, J. A. pp. 155, 292, 301, 337.
DAVIDS, N. p. 324.
DEB, S. p. 309.
DIEMINGER, W. p. 340.
EHMERT, ALFRED p. 189.
ELLISON, T. H. p. 14.
FÖRSTERLING, KARL p. 22.
GREENHOW, J. S. p. 282.
JACKSON, W. p. 160.
KASEMIR, H. W. p. 32.
KENNY, P. J. p. 266.
LANDMARK, B. p. 254.
LIED, F. p. 362.
LITTLE, C. G. p. 356.
LOVELL, A. C. B. pp. 83, 308 (B).
MASSEY, H. S. W. pp. 1, 253.
MAXWELL, A. p. 356.
MINNIS, C. M. p. 261.
NOLAN, P. J. p. 266.
NORTHOVER, F. H. pp. 106, 200.
O'DONNELL, G. A. p. 201.
OSBORNE, B. W. p. 66.
PAETZOLD, H. K. p. 183.
PHILLIPS, G. J. p. 141.
PIGGOTT, W. R. p. 236.
RAWER, K. pp. 38, 51.
REGENER, E. p. 173.
RUBIN, SYLVAN p. 130.
SHINN, D. H. pp. 85, 200.
SMITH, F. G. p. 350.
SMITH-ROSE, R. L. pp. 83, 260 (B).
SPENCER-JONES, SIR HAROLD p. 259.
STRANZ, DIETRICH p. 79.
WHALE, H. A. p. 85, 200.
WHITEHEAD, J. D. p. 361.
WÜSTER, HANS-OTTO p. 22.
YATES, G. G. p. 272.

Journal of Atmospheric and Terrestrial Physics, 1951, Vol. 2, pp. 1 to 13. Pergamon Press Ltd., London

The negative ion concentration in the lower ionosphere

D. R. BATES and H. S. W. MASSEY, F.R.S.

University College, London

(Received 3 May 1951)

ABSTRACT

The processes leading to the formation and destruction of negative ions in the upper atmosphere are considered in detail. Even if the former are improbably rapid, and the latter are improbably slow, the day time equilibrium concentration of negative ions in the D and E regions does not seem to be sufficiently high to be readily reconcilable with the dynamo theory of the L and S magnetic variations. Thus on calculating the transverse conductivity of E by the usual method it is found to be such that the magnitude of the L current can apparently only be explained if the local tidal motion is 6000 times greater than at ground level; and the most recent work on the interpretation of radio observations on the movements of the ionosphere is opposed to an amplification of this order. There seems also to be little doubt that the transverse conductivity of D is not as great as is demanded by the theory in its present form.

The fade-out enhancement of the electron concentration in the lower ionosphere is briefly discussed. It is pointed out that the possibility that it arises from photo-detachment by Lyman (α) radiation from solar flares cannot be excluded.

VOL. 2 1952

1. Since both oxygen atoms and molecules can form stable negative ions by attachment of electrons the question arises as to the magnitude of the negative ion concentration at any point in the ionosphere. This is usually specified in terms of a quantity λ, defined as the ratio of the concentration of negative ions to that of electrons. The value of λ is important for the interpretation of many phenomena. Thus the effective rate of recombination of electrons may depend on it; and so may the conductivity of the ionized layers.

In an exploratory paper in which the rates of various reactions occurring in the upper air were discussed, MASSEY [1] estimated that λ might be as large as 100 in the E region; but this early work made no allowance for photo-detachment. A subsequent investigation by BATES and MASSEY [2] showed that, in daytime, photo-detachment is actually so rapid that λ must be less than unity in all the main layers. In the lower ionized region, usually referred to as D, λ may be considerably greater than unity but BATES and SEATON [3] and MASSEY [4] estimated that, even at the 75 km level, it probably does not exceed 50 during the daytime. Apart from its importance in determining the properties of the ionization in reflecting and absorbing radio waves the value of λ in this region has become of great interest for the dynamo theory of the lunar and solar quiet-day magnetic variations. There is no insuperable obstacle to obtaining sufficient conductivity from E to meet the general requirements of this theory provided that in it λ can be taken as about 10 or greater, for too large an amplification of the tidal motion above that at ground level would not then be required. However, if λ is less than unity the position is not nearly so satisfactory. Furthermore, there are difficulties in the interpretation of the relations between the lunar and solar variations which apparently can be avoided only if the conductivity of the D region is greater than that of the E region [5] which requires that λ should be very large indeed throughout the lower region. The inadequacy of the estimated values of λ has not been considered as conclusive hitherto, it being argued that they

are too uncertain. The purpose of this paper is to reduce the uncertainty by a detailed examination of the possibilities, and to show that it is most unlikely that the true values of λ are sufficient. It will in fact be demonstrated that the expected transverse conductivity of the ionosphere is so much smaller than that required by the dynamo theory that the discrepancy must be regarded as of major significance. In general we will discuss more specifically the L variation (as for it there is at present no alternative to the dynamo theory) but our conclusions apply also to the dynamo theory of the S variation.

2. The Dynamo Theory

2. It is generally supposed that the lunar and solar quiet day magnetic variations arise from the magnetic fields of current systems set up in the ionosphere by a dynamo action. The lunar and solar tides produce motion of the conducting ionized layers across the Earth's magnetic field so inducing currents which flow in directions normal to the tidal flow and to the local magnetic field. The magnitude of the induced current at any point is proportional to the conductivity there and to the velocity of tidal flow.

Table 1. Values of λ and l_L in E region consistent with dynamo theory of lunar magnetic variations

λ	0	5	10	20
l_L	5×10^3	6×10^2	3×10^2	2×10^2

An important point arises when it is recognized that a conductor is rendered anisotropic by the presence of a magnetic field. The conductivity, $K^{\parallel}$, for flow parallel to the field is unaffected by the field but the transverse conductivity, $K^{\perp}$, for flow perpendicular to the field is reduced [6]. It has been generally assumed therefore that it is the lower transverse conductivity which determines the magnitude of the dynamo currents.

A detailed analysis of the L and S current systems, presumed responsible for the respective lunar and solar variations, has been carried out by Chapman and other [6]. In particular Chapman has found that to give the observed lunar magnetic variations at sunspot maximum the conductivity, integrated throughout the effective region must be $(2{\cdot}5\times10^{-5}\,l_L^{-1})$ e.m.u., l_L being the factor by which the mean speed of the lunar tide throughout the region exceeds that at ground level.

The conductivity required for the S current system is approximately the same except that l_L^{-1} is replaced by l_S^{-1} where l_S is the amplification factor for the solar tide. It may be noted that the amplitudes of the lunar and solar semi-diurnal pressure variations at ground level in equatorial regions are respectively about 8×10^{-5}, and $1{\cdot}2\times10^{-3}$ of an atmosphere. Proposed amplification factors whose reciprocals approach these figures cannot readily be accepted.

The transverse conductivity of the E region was calculated in 1945 by Cowling [7]. The values of l_L required for different assumed values of λ are given in Table 1 which is taken from his paper.

Evidence of a large amplification had already been obtained in 1939 through the remarkable investigations of Appleton and Weekes [8] on the lunar variation of the height of maximum ionization in the E region. Interpreted on the assumption that this height is determined solely by the pressure their results gave a relative lunar pressure variation some 6000 times as great as that at ground level. This would be sufficient to give the required lunar dynamo currents even if λ were not

very large. But certain difficulties remain. One is that the effect observed by APPLETON and WEEKES is 180° out of phase with that required by the dynamo theory. Another is that the interpretation of the observed height variations in terms of pressure variations is open to considerable doubt. Thus MARTYN [5] has pointed out that the electric fields driving ions and electrons across the horizontal component of the Earth's magnetic field will influence the vertical distribution of ionization and has shown that many previously unexplained variations of the distribution may be understood in this way. From a comparison of the lunar variation in the height of E region maximum electron density at Canberra with that at Cambridge he has indeed found that, whereas tidal pressure variations are inadequate to account for the observed effects, introduction of vertical motion of charged particles due to electrodynamic forces does provide a satisfactory explanation. Again WEEKES* has pointed out that analyses of the radio observations on ionospheric winds have given no indication of a component due to lunar tidal action though this would have been detected if the amplification exceeded a limit some 30 times smaller than the value originally suggested.

Table 2. Assumed atmospheric properties

Altitude km	*Temperature* deg K	*Total particle concentration* per cm³
50	290	$2{\cdot}3 \times 10^{16}$
75	205	$9{\cdot}5 \times 10^{14}$
90	245	$8{\cdot}7 \times 10^{13}$
100	270	$2{\cdot}3 \times 10^{13}$
110	305	$7{\cdot}4 \times 10^{12}$
120	360	$2{\cdot}7 \times 10^{12}$
150	(550)	$(2{\cdot}8 \times 10^{11})$
200	(900)	$(2{\cdot}8 \times 10^{10})$
250	(1200)	$(5{\cdot}6 \times 10^{9})$
300	(1500)	$(1{\cdot}8 \times 10^{9})$

Since COWLING's investigation, the density and temperature distribution in the upper atmosphere has been determined more accurately from rocket and other measurements [9]. We have repeated his calculations using the new data (which is summarized in Table 2) and the best available values for the ionic and electronic mobilities, but as shown in the following section the revision does not improve the position.

Apart from the problem of the magnitude of the L variations it has been pointed out by MARTYN [5], on the basis of his interpretation of the vertical distortion of the layers in terms of electrodynamic forces, that the phase of the lunar tide in E and F_1, is in the opposite sense to that required by the dynamo theory. As the temporal variations of F_2 are not easy to reconcile with the observed magnetic effects, and as frictional effects of various kinds are likely to reduce the amplitude of tidal flow in it. MARTYN was led to ignore this region (in spite of its considerable conductivity *c.f.* Figure 1) and to ascribe the in-phase component of the tide to dynamo currents in the D region. The discovery by MCNISH [10] that solar flares enhance the S variations was taken by MARTYN as evidence that the associated current sheet is also partially located in D since the ionization there is known to be increased**. Furthermore there are marked differences between the S and L quiet day magnetic variations as regards their dependence on the seasonal and sunspot cycles. Whereas the S variation

* Private communication.

** It cannot be claimed that the flare evidence *proves* that a substantial part of the S dynamo current flows near the 80 km level; for Mr. K. WEEKES kindly informs us that recent work at Cambridge indicates that (contrary to early belief) the augmentation of the ionization is more pronounced at greater altitudes.

VOL. 2 1952

exhibit little seasonal influence but a marked dependence on the sunspot cycle, the reverse is true of the L variations. To explain these differences as well as the phase relations MARTYN supposes that both the solar and lunar dynamo currents include contributions from the D as well as the E and F_1 regions so that, if S and L denote the respective total currents,

$$S = S_D + S_{EF_1}, \tag{1}$$

$$L = L_D - L_{EF_1}, \tag{2}$$

where the suffices D and EF_1 indicate the contributions from the respective regions. He then shows that the relative magnitudes of the solar and lunar variations as well as their dependence on season and sunspot cycle are obtained if the following assumptions are made:

ratio of the integral conductivity of D to that of $EF_1 = 1{\cdot}41$,
ratio of lunar to solar tidal velocity $= 0{\cdot}40$,
ratio of summer to winter conductivity of $D = 2{\cdot}7$,
ratio of summer to winter conductivity of $EF_1 = 2{\cdot}1$,
ratio of conductivity of D at sunspot maximum to that at sunspot minimum $=$ 1·43,

and if the ratio of the conductivity of EF_1 at sunspot maximum to that at sunspot minimum is taken to be 1·6, in agreement with the observed variation of electron density in these layers during the sunspot cycle.

From our point of view the most significant quantity is the first which assigns to the D region an integral conductivity rather greater than that of the E and F_1 regions. The grave difficulties which this encounters have already been pointed out by COWLING and BORGER [11] but it is possible to make the objections much more definite by showing that λ in D cannot even approach that required to provide sufficient transverse conductivity if this is calculated in the usual way. Before considering the matter further we shall discuss in detail the values of the conductivity and of λ throughout D and E to show how these conclusions (particularly the latter) have been reached.

3. The Transverse Conductivity for Ions and for Electrons at Different Atmospheric Levels

3. According to COWLING [7] the transverse conductivity due to particles of charge e and mass m at any particular level of the ionosphere is

$$K^{\perp} = \frac{n e^2}{m} \cdot \frac{\nu}{\nu^2 + \omega^2}, \tag{3}$$

where n is the number of particles per cm^3, ν is the collision frequency, and ω the gyro frequency.

Calculation of $K^{\perp}/n$ depends on a knowledge of ν which is given by

$$\nu = \sum n_r Q_r^d \bar{v}_r \tag{4}$$

where n_r is the concentration of neutral molecules of the r^{th} kind, Q_r^d is the momentum loss or diffusion cross section for collisions between these molecules and the electrons or ions concerned, and $\bar{v}_r$ is the mean velocity of relative motion. The effect

of collisions between ions and ions, or ions and electrons, is quite negligible except in the F_2 region [7].

Information about the value of ν_e, the collision frequency for electrons, may be obtained from experiments on the diffusion of very slow electrons in air and from measurements of the shift in the term values of the highly excited levels of the alkali metals due to addition of nitrogen and of oxygen to the emitting vapour*. In his calculations COWLING [7] assumed from the measurements of TOWNSEND and TIZARD [13] that

$$\sum n_r Q_r^d = 50 \text{ cm}^{-1}, \tag{5}**$$

at a pressure of 1 mbar and a temperature of 273° K. The more recent work of HUXLEY and ZAAZOU [14] using modern technique suggests

$$\sum n_r Q_r^d = 20 \text{ cm}^{-1}, \tag{6}**$$

for the same conditions, while observations on the pressure shift in the high series terms of the alkali spectra [15] give

$$\sum n_r Q_r^d = 5 \text{ cm}^{-1}. \tag{7}$$

To allow for the obvious incertainty two simple expressions for ν_e were adopted, $5\times10^{-9}(\sum n_r)$/sec and $5\times10^{-8}(\sum n_r)$/sec. At the pressure and temperature specified above the corresponding values of $\sum n_r Q_r^d$ are about 20 cm^{-1} and 200 cm^{-1} respectively. There can be little doubt therefore that ν_e is not underestimated. As will be seen both expressions when used in (3) give only a small contribution from the electrons to the transverse conductivity.

Precision measurements on the mobility of N_2^+ in molecular nitrogen (TYNDALL [16]) indicate that at temperature T the effective collision radius is $5{\cdot}9\times10^{-8}(300/T)^{\frac{1}{2}}$ cm. Following COWLING we used this value and adopted a mean molecular weight of 26 in calculating the collision frequency in the D and E regions. A more elaborate treatment is scarcely justified as the ions actually present have not yet been identified with certainty [3] and as the correction to be applied if they are not N_2^+ is quite insignificant. Thus JOHNSON [12] has given reasons for supposing that the effective collision radius is controlled mainly by the electric quadrupole moment of N_2, and is therefore almost independent of the species of ion. Again the extensive and accurate experimental investigations that have been carried out (TYNDALL [16]) show that, in harmony with JOHNSONs hypothesis, the mobility of a foreign ion in molecular nitrogen is a simple function of its mass. From the formula representing the laboratory results it is found that the expected mobility of O^+ (the lightest, and hence most mobile, ion that could be common) is only 1·25 times the measured mobility of N_2^+. At the gas densities prevailing in the D and E regions the gyrofrequency in the horizontal magnetic field of 0·4 gauss is much less than the collision frequency so the ionic conductivity would only be affected by approximately the same factor.

* Recently JOHNSON has attempted to calculate ν_e assuming that the main contribution arise from interaction of electrons with the quadrupole moment of N_2 [12]. As pointed out by Dr. JOHNSON in a private communication to us he has not been able to take account of the wave nature of the electrons which is decisive at the low electron velocities concerned.

** The experiments were conducted on electrons of energy considerably above thermal so that the results have to be extrapolated.

VOL. 2 1952

Figure 1 gives the values of the electronic $K_e^{\perp}/n(e)$ and the ionic $K_i^{\perp}/n(i)$ finally obtained. It will be noted that the maximum of the former occurs at about 85 km or less, and that of the latter occurs at some 150 km.

4. The Value of λ and the Conductivity of the D Region

4.1. Consideration of the equilibrium between the various processes leading to the formation and destruction of negative ions [2] shows that the ratio of their concentration to that of the free electrons is given by

$$\lambda = \eta\, n(X)/\{\varrho + \varkappa\, n(Y) + (1+\lambda)\,\alpha_i\, n(e)\}, \tag{8}$$

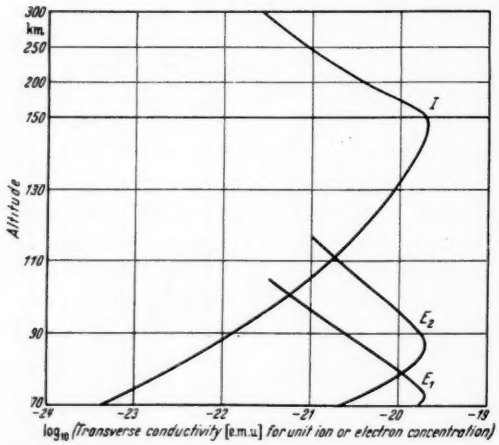


Fig. 1. Transverse conductivity of upper atmosphere I ionic contribution, E_1 and E_2 electronic contribution for assumed collision coefficients of 5×10^{-9} cm³/sec and 5×10^{-8} cm³/sec respectively.

where the n's are particle concentrations, η is the rate coefficient associated with radiative attachment

$$e + X \rightarrow X^- + h\nu, \tag{9}$$

or with any other attachment mechanism, ϱ is that associated with photodetachment by solar radiation

$$X^- + h\nu \rightarrow X + e, \tag{10}$$

$\varkappa$ is that associated with collision detachment

$$X^- + Y \rightarrow XY + e, \tag{11}$$

and α_i is that associated with two or three body ionic recombination

$$X^- + Z^+ \rightarrow X' + Z'', \tag{12}$$

$$X^- + Z^+ + M \rightarrow XZ + M. \tag{13}$$

Each of the reactants X, Y and Z is assumed to be a single species. This simplification is unimportant.

An attempt has recently been made [3], [4] to calculate λ from (8) using estimates of the various coefficients (which unfortunately are poorly determined). The values at the 75 km and 90 km levels were found to be about 20 and 2 respectively. From the viewpoint of collision theory these figures are still regarded as being of the most probable order of magnitude*. Clearly, however, it is desirable to reexamine the whole position to sec whether there is any possibility that λ is in fact much larger as is demanded by the geomagneticians.

In the D region molecular oxygen is by far the most abundant constituent having an appreciable positive affinity**. Attachment to it has been investigated experimentally by Bradbury, and by Healey and Kirkpatrick [17]. The former found the coefficient for low energy electrons to be about 5×10^{-12} cm³/sec at gas pressures

* They should actually be somewhat reduced as recent work (*c.f.* Table 2) indicates that the atmosphere is rather more tenuous than was believed a few years ago.

** It may be noted that the affinity of O_2 is about 1 ev [4].

of 3 mbars and above; the latter found it to be slightly smaller. BLOCH and BRADBURY [18] have suggested that capture takes place through a radiationless transition and that the excited ion formed in this way is then stabilized by having its excess energy removed. To account for the observed results the stabilization process must be extremely rapid: either the excited ion must emit a photon within about 10^{-9}secs, or its cross section towards de-activitation by a collision with an ordinary oxygen molecule must be some 10^3 times the gas kinetic. The first of these possibilities can be dismissed with almost complete confidence. Though even the second must be regarded with suspicion it is usually tentatively accepted for lack of an alternative. An immediate consequence of it is that when p the partial pressure of molecular oxygen, is sufficiently low, the coefficient may be expressed in the form

$$\eta = \eta_0 + kp. \tag{14}$$

BATES and MASSEY [19] have given reasons for supposing that η_0 does not exceed 10^{-14} cm^3/sec, that k is perhaps $1{\cdot}5\times10^{-12}$ cm^3/sec/mbar and that (14) may be used when p is less than about 1 mbar. The limit assigned to η_0 cannot be reached unless the life time of the excited ion towards photon emission is shorter than about 10^{-6}sec (which seems unlikely) or unless direct radiative attachment is of the so-called resonance type (and the measurements on elastic scattering give no indication of this phenomenon). The proposals regarding the magnitude of k and the range of validity of (14) are inter-related and are based on the assumption that the de-activation cross section can scarcely be much greater than the value mentioned above. They receive some support from the early experiments of BROSE [20] which suggest that pressure dependence sets in below about 2 mbars.

VOL. 2 1952

Setting η_0 equal to the limit quoted it can be seen that (14) gives $\eta(O_2|75\text{ km})$ to be about $1{\cdot}6\times10^{-14}$ cm^3/sec, and $\eta(O_2|90\text{ km})$ to be about $1{\cdot}1\times10^{-14}$ cm^3/sec, if k is $1{\cdot}5\times10^{-12}$cm^3/sec/mbar; and that it gives $\eta(O_2|75\text{ km})$ to be about 7×10^{-14}cm^3/sec, and $\eta(O_2|90\text{ km})$ to be about $1{\cdot}6\times10^{-14}$ cm^3/sec, if k is 10 times larger (an improbable eventuality). In view of their importance special attention should be paid to the status of these figures. We wish therefore to emphasize that they depend on the extrapolation of laboratory data; and that the observed coefficients at moderate pressures are so large that they cannot be accepted with confidence. Because of this doubt (which should be resolved by further experiments) and because of the rather high value adopted for η_0 the estimated coefficient may well be too large. It is of course conceivable (though most improbable) that some very efficient attachment mechanism has been overlooked, and that the prediction that the coefficients at low pressures are smaller than the measured 5×10^{-12} cm^3/sec is erroneous. There can be little doubt that only a radiative process could have a coefficient constant down to the pressures prevailing in the D region.

4.3. On the reasonable assumption that the transition matrix elements involved in photo detachment from O_2^- and from O^- are of the same order, $\varrho(O_2^-)$ has previously been taken to be about 5×10^{-1}/sec. However the action of radiation on molecular structures exhibits some peculiarities: for example photo-ionization of O_2 above the 12·2 ev potential apparently occurs far less readily [21] than would be expected from simple theory. It may therefore be that the true value $\varrho(O_2^-)$ is smaller than that originally adopted. To enable the factors entering to be appreciated easily it is useful to divide the continuum into a series of bands each of width 1 ev

and to write

$$\varrho(O_2^-) \simeq \sum_{\varepsilon=1\cdot5}^{\infty} I_\varepsilon A_\varepsilon \tag{15}$$

where I_ε is the incident flux of solar photons in the band centered at ε ev and A_ε is the mean detachment cross section (in cm^2), through this band. As a crude approximation we may take the Sun to emit as a 6000°K black body up to photon energies of almost 4 ev and as a 5000°K black body at higher photon energies. Hence on numerical substitution we have

$$\left.\begin{aligned}\varrho(O_2^-) \simeq 10^{17}\{&2\cdot8\,A_{1\cdot5}+1\cdot1\,A_{2\cdot5}+2\cdot5\times10^{-1}A_{3\cdot5}+\\&+1\cdot6\times10^{-2}A_{4\cdot5}+2\cdot2\times10^{-3}A_{5\cdot5}+\cdots\}/\text{sec}\,.\end{aligned}\right\} \tag{16}$$

In the case of atomic negative ions the cross section rises from zero at the spectral head to a maximum of usually at least 10^{-18}–10^{-17} cm^2 at an energy of order twice that required for detachment. Cancellation in the transition matrix element is not to be expected in the low energy region [22] and in any case it could only be effective over a narrow range. Clearly therefore (16) suggests that $\varrho(O_2^-)$ is unlikely to be very much less than the value quoted above—thus it is as great as $1\cdot7\times10^{-2}$/sec even if, to take a somewhat extreme example

$$\left.\begin{aligned}A_{1\cdot5}&=5\cdot4\times10^{-21}\,\text{cm}^2\\A_{2\cdot5}&=8\cdot2\times10^{-20}\,\text{cm}^2\\A_{3\cdot5}&=2\cdot3\times10^{-19}\,\text{cm}^2\\A_{4\cdot5}&=3\cdot9\times10^{-19}\,\text{cm}^2\\\text{and}\quad A_{5\cdot5}&=5\cdot0\times10^{-19}\,\text{cm}^2\ \text{(the maximum)}\end{aligned}\right\} \tag{17}$$

It is instructive to consider the oscillator strength f_c associated with the continuum. From a well known quantal formula we readily obtain the relation

$$f_c \simeq 9\times10^{15}\sum_{\varepsilon=1\cdot5}^{\infty} A_\varepsilon\,. \tag{18}$$

Inserting the values of A_ε just used we see that their contribution to f_ε is only about 0·01 and the contribution from the high energy region can scarcely be much greater. This result suggests that, far from being too large, the values assumed for A_ε are probably too small. To be sure the greater part of the total oscillator strength of the electron shell concerned might well be associated with discrete transitions. But, in this case, these would have to be taken into account in calculating the rate of destruction of the ions for they would almost certainly give rise to autodetachment. The THOMAS-KUHN sum rule thus makes it difficult to believe that $\varrho\,(O_2^-)$ is many orders smaller than the 5×10^{-1}/sec originally proposed—a value of 5×10^{-2}/sec is doubtless possible, one of 5×10^{-3}/sec is only barely acceptable.

4·4. Various combinations of the figures discussed were used in conjunction with the simplified form of (8) obtained by omitting the $\varkappa$ and α_i terms from the denominator. Table 3 gives the values of λ deduced together with (i) the corresponding values of $n(O_2^-)$ (it being assumed that $n(e)$ is $2\cdot5\times10^2$/cm^3 at the 75 km level, and is $1\cdot5\times10^4$/cm^3 at the 90 km level (*c.f.* [3]) and (ii) the ratios R of the CHAPMAN conductivity ($2\cdot5\times10^{-5}$ e.m.u.) to the calculated conductivity of a 10 km thick layer. For convenience in assessing the significance of the results we have affixed rating

Table 3. The transverse ionic conductivity in the D region

Altitude *km*	*Assumed attachment coefficient* $\eta(O_2)$ *cm³/sec*	*Assumed photo detachment coefficient* $\varrho(O_2^-)$ *cm³/sec*	*Deduced negative ion to electron ratio* λ	*Corresponding ion concentration* $n(O_2)$ *cm³/sec*	*Ratio R of* CHAPMAN *conductivity to calculated ion conductivity of a 10 km layer*
75	$1{\cdot}6 \times 10^{-14}$ (*A*)	5×10^{-1} (*A*)	6	2×10^{3}	6×10^{8}
	7×10^{-14} (*B*)		3×10^{1}	7×10^{3}	1×10^{8}
	5×10^{-12} (*C*)		2×10^{3}	5×10^{5}	2×10^{6}
	$1{\cdot}6 \times 10^{-14}$ (*A*)	5×10^{-2} (*B*)	6×10^{1}	2×10^{4}	6×10^{7}
	7×10^{-14} (*B*)		3×10^{2}	7×10^{4}	1×10^{7}
	5×10^{-12} (*C*)		2×10^{4}	5×10^{6}	2×10^{5}
	$1{\cdot}6 \times 10^{-14}$ (*A*)	5×10^{-3} (*C*)	6×10^{2}	2×10^{5}	6×10^{6}
	7×10^{-14} (*B*)		3×10^{3}	7×10^{5}	1×10^{6}
	5×10^{-12} (*C*)		2×10^{5}	5×10^{7}	2×10^{4}
90	$1{\cdot}1 \times 10^{-14}$ (*A*)	5×10^{-1} (*A*)	4×10^{-1}	6×10^{3}	2×10^{7}
	$1{\cdot}6 \times 10^{-14}$ (*B*)		6×10^{-1}	8×10^{3}	1×10^{7}
	5×10^{-12} (*C*)		2×10^{2}	3×10^{6}	4×10^{4}
	$1{\cdot}1 \times 10^{-14}$ (*A*)	5×10^{-2} (*B*)	4	6×10^{4}	2×10^{6}
	$1{\cdot}6 \times 10^{-14}$ (*B*)		6	8×10^{4}	1×10^{6}
	5×10^{-12} (*C*)		2×10^{3}	3×10^{7}	4×10^{3}
	$1{\cdot}1 \times 10^{-14}$ (*A*)	5×10^{-3} (*C*)	4×10^{1}	6×10^{5}	2×10^{7}
	$1{\cdot}6 \times 10^{-14}$ (*B*)		6×10^{1}	8×10^{5}	1×10^{5}
	5×10^{-12} (*C*)		2×10^{4}	3×10^{8}	4×10^{2}

VOL. 2 1952

letters *A*, *B* or *C* to the different coefficients employed. Those marked *A* are considered the best estimates that can be made on the evidence at present available (which is not to be taken as meaning that they are regarded as reliable); those marked *B* are thought to be much less likely to be correct; and those marked *C* are viewed with extreme scepticism.

The ratio *R*, which approximates to the amplification of the lunar tidal oscillation that would be required to explain the magnetic variations, is manifestly very large in all cases. It is clear that the situation at the 75 km level is most unpromising. At the 90 km level the value of *R* falls to 4×10^2 in one instance. An amplification of this order is perhaps conceivable; but in view of the fact that both of the coefficients concerned are of rating *C*, little credence can be given to their combination. It must be born in mind too that the use of a high $\eta(O_2)$ and of a low $\varrho(O_2^-)$ is only justified if electron capture takes place without the emission of radiation*. For if radiation were emitted the attachment process would have an inverse photo detachment process and therefore $\eta(O_2)$ and $\varrho(O_2^-)$ would be thermodynamically related. Elementary calculations show that in this event λ would be of the order of unity or less throughout the atmospheric region under discussion.

4·5. Quite apart from photo detachment λ is restricted by the loss of negative ions in recombination. The two-body contribution (12) to α_i is unknown but fortunately the three body contribution (13) can be estimated with fair reliability. By using the well-established theory of THOMSON [22] to extrapolate the experimental data

* See the remark at the end of § 4.2 on the improbability that any process but a radiative one could have a coefficient which is a high in the *D* layer as the measured laboratory value.

of SAYERS and GARDNER [24] to low pressures its value was found to be about 1×10^{-9} cm^3/sec at the 75 km level and 5×10^{-11} cm^3/sec at the 90 km level. Hence from (8) we see that even if the attachment coefficient is the maximum conceivable, and even if the negative ions disappear solely by recombination, λ at 75 km could only be 6×10^4, and at 90 km only 1×10^4, both of which figures are smaller than the upper limits previously deduced. It may be noted also that $(\lambda n(e))^2\alpha_i$ would be at least 2×10^5/cm^3/sec and 1×10^6/cm^3/sec respectively at the two levels concerned. An extremely high ionization rate would therefore be required to maintain the layer*.

Again, an O_2^- ion can be destroyed directly by collision detachment,

$$O_2^- + O \rightarrow O_3 + e, \tag{19}$$

or its excess electron can be transferred to an atom from which it would certainly be rapidly removed by the action of solar radiation,

$$O_2^- + O \rightarrow O_2 + O^-, \tag{20}$$

$$O_2^- + H \rightarrow O_2 + H^-, \tag{21}**$$

$$O_2^- + H \rightarrow OH + O^-. \tag{22}$$

The work of BATES and NICOLET [9] suggests that the concentration of O is some 10^{11}/cm^3 at both 75 km and 90 km, and that that of H is some 10^{10}/cm^3 at the former altitude and 10^9/cm^3 at the latter. If these concentrations are correct, and if the coefficients $\varkappa$ of the above process satisfy the following conditions.

$$\left.\begin{aligned} &75 \text{ km or } 90 \text{ km, process (19) or (20)}, && \varkappa \geq 5\times10^{-14} \text{ cm}^3/\text{sec} \\ &75 \text{ km, process (21) or (22)}, && \varkappa \geq 5\times10^{-13} \text{ cm}^3/\text{sec} \\ &90 \text{ km, process (21) or (22)}, && \varkappa \geq 5\times10^{-12} \text{ cm}^3/\text{sec} \end{aligned}\right\} \tag{23}$$

there would be a restriction on λ as effective as that which is provided by photo detachment with a value for $\eta(O_2^-)$ of 5×10^{-3}/sec. It is not known if any of the demands of (23) are actually met. But as they are not severe we have yet another reason for supposing that we have not grossly overestimated the rate of destruction of negative ions. Furthermore, if attachment is rapid, and if collision detachment does not occur, it is difficult to understand why the electrons do not disappear almost completely within less than a second of sunset.

It might be suggested that, as water vapour molecules are highly polar, their presence in the upper atmosphere might profoundly influence the whole situation. However, their abundance relative to oxygen is probably less than 10^{-4} at altitudes of 75 km and less than 10^{-6} at altitudes of 90 km [9]. Such factors could not be counterbalanced by any conceivable attachment coefficient; and in any case the supposition that attachment to H_2O is abnormally effective is contrary to the laboratory evidence available [4]. The other minor atmospheric constituents can similarly be dismissed as unimportant in the present connection.

Complex ions may exist and if their binding energy is large compared with the mean energy of thermal motion they may even be more numerous than simple ions.

* For comparison it may be recalled that the measured ionization rate at the maximum of the *E* ayer is only about 2×10^2/cm^3/sec.

** A decision cannot yet be made as to whether (21) is exothermic as the electron affinity of O_2 is not known sufficiently accurately.

Since however their formation is ultimately controlled by the attachment processes already discussed, and since they are subject to at least one of the destructive processes (recombination) they cannot affect the limit to λ that has been deduced.

We are forced to conclude that, if the relevant conductivity per ion and per electron is given by (3), the value of λ is very much too small in the D region for dynamo currents to make any appreciable contribution to the lunar magnetic variations.

5. The Value of λ and the Conductivity of the E Layer

Atomic oxygen is abundant at the altitudes of the E layer and because of its relatively high electron affinity (2·2 ev) it might be expected to be the principal negative ion forming constituent. The processes leading to the production and the destruction of O^- were therefore examined in some detail in earlier papers [2]. It was shown that the contribution to λ from this particular ion cannot *even in the most favourable circumstances* exceed $7\times10^{-14}\,n(O)$ which is certainly smaller than unity since $n(O)$ the concentration of oxygen atoms, is at most some $10^{13}/cm^3$. The discussion originally given need not be repeated here. It has been subjected to the closest scrutiny but no reason has been found for supposing that it requires modification.

Admittedly account was not taken of excited metastable O^- ions, for whose existence there is now considerable experimental evidence [4]. However, under upper atmospheric conditions, such ions could only be formed by radiative attachment which must be very slow as the energy of the emitted photons is extremely small; and from thermodynamic arguments it may be proven that photo-detachment is sufficiently rapid to prevent their equilibrium concentration from being appreciable.

VOL. 2 1952

The presence of molecular oxygen can scarcely affect the position to any material extent. At the 100 km level its partial pressure is probably only some 10^{-4} mbars or less [9]. It is therefore difficult to believe that appreciable attachment can take place other than by a radiative process: and if this is accepted calculations based on thermodynamics can be used at once to show that λ must be below unity. Owing to photo-dissociation the molecular oxygen concentration falls off rapidly at altitudes above 100 km [25]. Consequently the argument just presented can be applied with even greater confidence at altitudes of 110 km and greater. It will be noted that strong supporting evidence is provided by the fact that the dependence of the peak electron density $n_0(e)$ on the solar zenith angle χ is almost as predicted by the Chapman theory [26]. This follows immediately from the near-equilibrium equation

$$n_0(e) = \left\{\frac{i_0\cos\chi}{(1+\lambda)(\alpha_e+\lambda\,\alpha_i)}\right\}^{\frac{1}{2}}, \tag{23}$$

where i_0 is the peak ionization rate at the sub-solar point, α_e is the electronic recombination coefficient, and the other symbols have their usual significance [2]. For, in order that $n_0(e)$ be approximately proportional to $(\cos\chi)^{\frac{1}{2}}$ (as is required), it is clearly necessary that λ be pressure independent (which is inconceivable) or that it be small.

In view of the importance of the issue it is perhaps worth considering once again the situation at the 100 km level to see whether λ could be large there even if non-radiative attachment were dominant. As an extreme example suppose that the de-activation cross section mentioned in § 4.1 is 10^4 times the gas kinetic and that stabilization by photon emission does not occur. In this event $\eta(O_2|100\,km)$ would

be $1{\cdot}5\times10^{-15}$ cm^3/sec. The minimum acceptable value of $\varrho(O_2^-)$ is 5×10^{-3}/sec.* Taking the concentration of molecular oxygen to be 10^{13}/cm^3 and using (8) we find once more that λ is only unity or less. To be sure a high λ would result if $\eta(O_2)$ remained at the laboratory value down to a pressure of 10^{-4} mbars; but a mechanism with such a coefficient is unknown to collision theory.

Taking λ to be effectively zero, and assuming E to approximate to a Chapman layer of scale height 10 kms and maximum electron density $1{\cdot}5\times10^5$/cm^3 at 120 km, it may be shown that the integrated conductivity up to 150 km is about 4×10^{-9} e.m.u. To obtain the required lunar dynamo current the amplification factor would therefore have to average as much as 6000 throughout the effective thickness of the region. Although it cannot be said with complete certainty that this is an aerodynamic impossibility, a very special temperature distribution in the E region would be required to give it. In any case according to Weeks (*c.f.* 2) the measurements on ionospheric winds indicate that the actual amplication factor is at most 200; so the calculated conductivity would appear to be at least 30 times too small. The possible errors in the atomic parameters used in the evaluation of (3) are quite insufficient to cause such a discrepancy.

6. Concluding Discussion on Dynamo Theory

It seems quite clear that the negative ion concentration in D is wholly inadequate to provide dynamo currents of significant magnitude to contribute to the L and S quiet day magnetic variations if it is assumed as usual that it is the transverse conductivity given by (3) which is effective. In this connection it is of interest to note that Maple, Bowen and Singer [27] have obtained direct evidence, from rocket measurement of magnetic fields in the atmosphere near the geomagnetic equator, of a current system of considerable strength near the 100 km level but none of any appreciable currents below. Apart altogether from the difficulty of explaining the different seasonal and sunspot dependence of the L and S variations the supposition that the currents flow solely in E is open to objection in that the conductivity of this region, calculated in the same way, is so low that amplification of the lunar tide by about 6000, throughout an altitude range of some 30 km, would be required to give the necessary magnitude of the lunar dynamo current. Although it must be remembered that Appleton and Weekes found that over Cambridge the height variation of the E region maximum exhibits a lunar variation which could be interpreted as implying an amplification of just this value, the phase of the variation is opposite to that required, the effect is much smaller above Canberra and bears such phase relation to that at Cambridge as to make it likely that the height variations are due to vertical drift under electrodynamic forces, rather than to tidal pressure variations; and there are other objections mentioned in the main text (§ 2). Taken as a whole the position must be regarded as very unsatisfactory. No alternative suggestions can be made here but it is perhaps pertinent to note that there is still doubt as to whether the formulae (3) is really the correct one to use [28]. The possibility that the effect of the Earth's magnetic field is cancelled by polarization has been raised [29] but Cowling and Borger [30] have given reasons why this cancellation is only

* It may be remarked that, as the concentration of atomic oxygen at 100 km is of order 10^{13}/cm^3, process (19) or (20) would destroy negative ions as rapidly as the photo-detachment here envisaged if the coefficient associated with either were even 5×10^{-16} cm^3/sec.

partial. It is conceivable that local fluctuations in ion and electron density may assist in reducing the influence of the magnetic field but it is difficult at present to assess the importance of the effects.

7. The Role of Negative Ions in Fade-out Phenomena

Radio fade-outs arise from sudden enhancement of the electron density in the D region. They are associated with an incidence of solar flares on the Sun which may lead, among other effects, to enhanced emission of Lyman (α) radiation. As has been pointed out earlier [3] the very rapid recovery sometimes observed after the fade-out [31] can most readily be attributed to loss of electrons by attachment. The implication that λ is reduced during the fade-out might seem difficult to accept since the intensity of the visible and near ultraviolet part of the solar spectrum is not increased appreciably by a flare. However it is not necessarily correct to assume that the far ultra-violet part gives a negligible contribution to the photo detachment rate. Suppose, for example, that in the region of Lyman (α) the photo detachment cross section is about 5×10^{-17} cm^2 and that during a fade-out the emission in a 3 Å band centred on this line is 10^5 times that from a 6000° K black body. It may be shown that in these hypothetical circumstances $\varrho(O_2^-)$ would be augmented by about 1×10^{-1}/sec which *may* be an appreciable fraction of its normal value. Consequently there is at least the *possibility* that a flare can cause λ to decrease. Although the estimated value of λ is far too small to provide the transverse conductivity required in D by the present dynamo theory it is quite large enough for photo-detachment to cause the observed enhancement of electron density during a flare.

Acknowledgements—We wish to thank Mr. W. R. Piggot and Mr. K. Weekes for stimulating discussions.

VOL. 2 1952

References

[1] Massey, H. S. W.; Proc. Roy. Soc. A 1937 **163** 311. [2] Bates, D. R. and Massey, H. S. W.; Proc. Roy. Soc. A 1946 **187** 261, *ibid.* 1947 **190** 1. [3] Bates, D. R. and Seaton, M. J.; Proc. Phys. Soc. B 1950 **63** 129. [4] Massey, H. S. W.; Negative Ions, (2nd Edition) Cambridge University Press, 1950. [5] Martyn, D. F.; Proc. Roy. Soc. A 1948 **194** 445. [6] Chapman, S., and Bartels, J.; Geomagnetism, Oxford University Press, 1940. [7] Cowling, T. G.; Proc. Roy. Soc. A 1945 **183** 453. [8] Appleton, E. V. and Weekes, K.; Proc. Roy Soc. A 1939 **171** 171. [9] Bates, D. R. and Nicolet, M.; J. Geophys. Res. 1950 **55** 301. [10] McNish, A. G.; Terr. Magn. Atmos. Elect. 1937 **42** 109. [11] Cowling, T. G. and Borger, R.; Nature, Lond. 1948 **161** 515. [12] Johnson, M. H.; Phys. Rev. 1950 **80** 101. [13] Townsend, J. S. and Tizard, H. T.; Proc. Roy. Soc. A 1913 **88** 336. [14] Huxley, L. H. G. and Zaazou, A. A.; Proc. Roy. Soc. A 1949 **196** 402. [15] Fuchtbauer, F., Schulz, P. and Brandt, A.F.; Z. Phys. 1934 **90** 403. [16] Tyndall, A. M.; Positive Ions, Camb. Univ. Press, 1938. [17] Bradbury, N. E.; Phys. Rev. 1933 **44** 883; and Healey, R. H. and Read, J. W.; Behaviour of Slow Electrons in Gases, Amalgamated Wireless Ltd, Australasia, 1941. [18] Bloch, F. and Bradbury, N. E.; Phys. Rev. 1935 **48** 689. [19] Bates, D. R. and Massey, H. S. W.; Phil. Trans. Roy. Soc. A 1943 **239** 269. [20] Brose, H.; Phil. Mag. 1925 **50** 536. [21] Chapman, S. and Price, W. C.; Rep. Progr. Phys. 1936 **3** 42. [22] Bates, D. R.; Mon. Not. Roy. Astr. Soc. 1946 **106** 432. [23] Thomson, J. J.; Phil. Mag. 1924 **47** 337. [24] Sayers, J.; Proc. Roy. Soc. A 1938 **169** 83. Gardner, M. E.; Phys. Rev. 1938 **53** 75. [25] Penndorf, R.; J. Geophys. Res. 1949 **54** 7. [26] Chapman, S.; Proc. Phys. Soc. 1931 **43** 26 and 433. [27] Maple, E., Bowen, W. A. and Singer, S. F.; J. Geophys. Res. 1950 **55** 115 and *ibid.* 1951, **56**, 265. [28] Schluter, A.; Z. Naturforschung a 1950 **5** 72. [29] Martyn, D. F.; Nature, Lond. 1948 **162** 142. [30] Cowling, T. G. and Borger, R.; Nature, Lond. 1948 **162** 143. [31] Bracewell, R. N. and Straker, T. W.; Mon. Not. Roy. Astr. Soc. 1949 **109** 28.

Journal of Atmospheric and Terrestrial Physics, 1951, Vol. 2, pp. 14 to 21. Pergamon Press Ltd., London

The propagation of sound waves through a medium with very small random variations in refractive index

T. H. Ellison

Clare College, Cambridge

(Received 17 August 1951)

ABSTRACT

Mathematical techniques, developed for the theory of turbulence, are applied to the study of the propagation of waves in a medium in which the refractive index varies slightly with position in a random manner. Both a diffraction and a ray theory are used to relate the statistical properties of the wave to those of the refractive index field, and the conditions for the validity of the ray theory are exposed. A particular result is, that, for a wave travelling in a statistically homogeneous medium where the scale of the variations in refractive index is suitably large compared with the wave length but small compared with the path length, the variations in intensity produced by the medium are proportional to the cube of the path length for short paths, but directly proportional to it for long paths.

Introduction

There are many problems in which it is necessary to consider waves travelling through a medium in which there are very small random variations of refractive index. In particular, the refractive indices of the atmosphere for sound, light and radio-waves depend on the temperature and humidity, so that fluctuations of temperature and humidity produced by turbulence cause scattering of the waves and fluctuations of wave intensity. That measurements of these fluctuations could be used as a source of information about turbulence was suggested some years ago by Dahl and Devik [1] and by Bovsheverov and Krassilnikov [2] and since then there has been a steady increase of interest in this aspect of the subject. There have been measurements of sound propagation by Krassilnikov [3], Knudsen [4], Krassilnikov and Ivanhov-Shits [5] and by Richardson [6] and measurements of the propagation of electromagnetic waves by Liedentopf and Wisshak [7], Megaw [8] and by Butler [9]. However there have been very few theoretical attempts, the most notable being by Krassilnikov [10], to relate the statistical properties of the scattering to the statistical properties of the refractive index field.

In this paper the properties of the refractive index field are assumed known and deductions made about the scattering. The rather more difficult and complicated problem of obtaining information about the refractive index field from observations of the scattering is not considered here.

It is simple to perform the analysis of this problem using the methods of geometrical optics, but while such a treatment should be adequate for the case of the propagation of light over short paths in the atmosphere, it will be expected to be in error when the path length is such that the size of the first Fresnel zone in the parts of the refractive index field most distant from the receiver is no longer small compared with a distance in which there is appreciable variation in refractive index. This is frequently the case for radio and sound waves.

Unfortunately electro-magnetic waves do not satisfy the same wave equation as sound waves, and the treatments required in the two cases, though very closely parallel, are different in detail. For the sake of clarity, therefore, we shall confine ourselves here to the study of the propagation of sound waves in a stationary medium in which there are small variations in elasticity. The first approximation of the diffraction theory is calculated for the case of an incident plane wave, and is compared with the ray theory. The comparison is only made for one statistical quantity related to the scattering (the lateral spectrum for the fluctuation in intensity), but the extension of the theory to others, such as the mean square angular deviation of a ray, or the phase fluctuation, presents no difficulty.

Problems in which the incident wave is not plane are much more readily treated on the ray theory than on diffraction theory, so it is desirable that there should be a clear understanding of the limits of the ray theory and the way in which it breaks down.

The Wave Equation

We consider the equation

$$\triangle^2 \Psi - \frac{1-F(x,\,y,\,z)}{c^2}\,\frac{\partial^2 \Psi}{\partial t^2} = 0. \tag{1}$$

VOL. 2 1952

If $F(x, y, z)$ is very small, this is the wave equation for a medium in which there are very small variations of refractive index of magnitude $-\frac{1}{2}F(x, y, z)$.

Waves of only one frequency $(\omega/2\pi)$ will be considered so

$$\Psi = \Re\, \psi(x,\,y,\,z)\, e^{-i\omega t} \tag{2}$$

where $\Re$ denotes "the real part of."

ψ will then be a complex function of position and will have to satisfy

$$\nabla^2 \psi + n^2 \psi = n^2 F(x,\,y,\,z). \tag{3}$$

This can be transformed (see, for example, Mott and Massey [11]) into

$$\psi(x,\,y,\,z) = G - \frac{n^2}{4\pi} \iiint \frac{e^{in|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}\, F(x',\,y',\,z')\, \psi(x',\,y',\,z')\, dx'\, dy'\, dz' \tag{4}$$

where r is the vector whose components are (x, y, z), and G is a solution of $\nabla^2\psi + n^2\psi = 0$.

(4) can be solved by the method of successive approximation. Since the correction term at any order of the approximation corresponds physically to waves which have been scattered that same number of times, it seems likely, when $F(x, y, z)$ is very small, that the first few terms of the approximation will represent ψ very accurately. However, since the mean square error at each stage in the approximation depends on higher moments of the error at the previous stage, the author has not been able to fix rigorously the maximum error at any stage.

We consider an incident plane wave given by

$$\Psi = \Re\, e^{i(nz-\omega t)} = \cos(nz - \omega t)$$

so we take

$$G = e^{inz} \tag{5}$$

THE FIRST APPROXIMATION FOR ψ

From (4) and (5) we have

$$\psi = e^{inz} - \frac{n^2}{4\pi}\iiint \frac{e^{in|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} F(x', y', z')\,\psi(x', y', z')\,dx'\,dy'\,dz' \tag{6}$$

so the equation for the first approximation to ψ is

$$\psi_1 = e^{inz} - \frac{n^2}{4\pi}\iiint \frac{e^{in|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} F(x', y', z')\,e^{inz'}\,dx'\,dy'\,dz'\,. \tag{7}$$

From this point we shall assume, for mathematical convenience, that $F(x, y, z)$ is zero for $|x|$ or $|y| > L/2$ where L is some arbitary large length. L will not appear in the results of the calculation, so this procedure does not limit the application of the theory to physical problems. It enables us to take the two-dimensional Fourier transform of $F(x, y, z)$ in the form

$$\varphi(k_1, k_2, z) = \frac{1}{2\pi L}\int_{-\frac{L}{2}}^{\frac{L}{2}}\int_{-\frac{L}{2}}^{\frac{L}{2}} F(x, y, z)\,e^{+i(k_1 x + k_2 y)}\,dx\,dy\,. \tag{8}$$

The inverse of this is

$$F(x, y, z) = \frac{L}{2\pi}\iint_{\infty} \varphi(k_1, k_2, z)\,e^{-i(k_1 x + k_2 y)}\,dk_1\,dk_2\,. \tag{9}$$

Combining (7) and (9) we have

$$\psi_1 = e^{inz} - \frac{n^2 L}{8\pi^2}\iiiint\iint \frac{e^{in|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} e^{-i(k_1 x' + k_2 y' - nz')}\varphi(k_1, k_2, z')\,dk_1\,dk_2\,dx'\,dy'\,dz', \tag{10}$$

$$= e^{inz} - \frac{n^2 L}{8\pi^2} I_1(x, y, z) \quad \text{(say)}. \tag{11}$$

The evaluation of I_1 is fairly straightforward; the result is

$$I_1 = -2\pi i \iint (n^2 - k^2)^{-\frac{1}{2}}\, e^{+i[k_1 x + k_2 y + nz' + (n^2 + k^2)^{\frac{1}{2}}|z - z'|]}\,\varphi(k_1, k_2, z')\,dz'\,dk_1\,dk_2,$$

where $(n^2 - k^2)^{\frac{1}{2}}$ is to be taken as $+i(k^2 - n^2)^{\frac{1}{2}}$ when $k > n$.

So

$$\psi_1 = e^{inz} + \frac{i n^2 L}{4\pi}\iiint (n^2 - k^2)^{-\frac{1}{2}}\, e^{i[k_1 x + k_2 y + nz' + (n^2 - k^2)^{\frac{1}{2}}|z - z'|]}\,\varphi(k_1, k_2, z')\,dz'\,dk_1\,dk_2\,. \tag{12}$$

THE LATERAL SPECTRUM FOR THE FLUCTUATION IN INTENSITY

A typical quantity which might be measured in an investigation into the statistical properties of the wave after passing through the scattering layer is the covariance between the fluctuations of the wave intensities at pairs of points separated by a distance (ξ, η) in some plane perpendicular to the z axis, say $z = z_0$. If $E(x, y, z_0)$ is the relative fluctuation in wave intensity (*i.e.* the fractional difference between the instantaneous value of the flux of energy in the z direction per unit area and its mean value), this covariance may be defined by

$$R(\xi, \eta) = \frac{1}{L^2}\int_{-\frac{L}{2}}^{\frac{L}{2}}\int_{-\frac{L}{2}}^{\frac{L}{2}} \overline{E(x, y, z_0)\,E(x + \xi, y + \eta, z_0)}\,dx\,dy\,, \tag{13}$$

where L has the same meaning as in the previous section, and the overbear denotes a mean over many trials. It can be shown, that as L tends to infinity

$$R(\xi,\eta) \to \frac{1}{L^2}\int_{\infty}^{\infty}\int_{\infty}^{\infty} \overline{E(x,y,z_0)\,E(x+\xi,y+\eta,z_0)}\,dx\,dy. \tag{14}$$

In theoretical work it is simpler to use the two-dimensional Fourier transform of $R(\xi,\eta)$ with respect to ξ and η rather than $R(\xi,\eta)$ itself. This is called the lateral spectrum for the fluctuation in intensity and will be denoted by $I(k_1,k_2)$

$$I(k_1,k_2) = \frac{1}{(2\pi)^2}\iint R(\xi,\eta)\,e^{+i(k_1\xi+k_2\eta)}\,d\xi\,d\eta. \tag{15}$$

It is well known and easily proved that if

$$E(x,y,z_0) = \frac{L}{2\pi}\iint A(k_1,k_2)\,e^{i(k_1x+k_2y)}\,dk_1\,dk_2$$

then

$$I(k_1,k_2) = \overline{A(k_1,k_2)\,A(-k_1,-k_2)}. \tag{16}$$

We now make use of (12) to calculate $E(x,y,z_0)$ and hence $I(k_1,k_2)$.
From (2) and (12) we have

$$\left.\begin{aligned}
&\Psi(x,y,z,t) = \cos(nz-\omega t) + \\
&+\frac{in^2L}{8\pi}\iiint (n^2-k^2)^{-\frac{1}{2}}\varphi(k_1,k_2,z')\,e^{+i(k_1x+k_2y+nz'+(n^2-k^2)^{\frac{1}{2}}|z-z'-\omega t)}\,dk_1\,dk_2\,dz' - \\
&-\frac{in^2L}{8\pi}\iiint (n^2-k^2)^{-\frac{1}{2}}\varphi^*(k_1,k_2,z')\,e^{-i(k_1x+k_2y+nz'+(n^2-k^2)^{\frac{1}{2}}|z-z'|-\omega t)}\,dk_1\,dt_2\,dz' \\
&\qquad\qquad \text{for } n>k \\
&=\cos(nz-\omega t) + \\
&+\frac{n^2L}{8\pi}\iiint (k^2-n^2)^{-\frac{1}{2}}\varphi(k_1,k_2,z')\,e^{-(k^2-n^2)^{\frac{1}{2}}|z-z'|+i(k_1x+k_2y+nz'-\omega t)}\,dk_1\,dk_2\,dz' + \\
&+\frac{n^2L}{8\pi}\iiint (k^2-n^2)^{-\frac{1}{2}}\varphi^*(k_1,k_2,z')\,e^{-(k^2-n^2)^{\frac{1}{2}}|z-z'|-i(k_1x+k_2y+nz'-\omega t)}\,dk_1\,dk_2\,dz' \\
&\qquad\qquad \text{for } n<k.
\end{aligned}\right\} \tag{17}$$

Since $F(x,y,z)$ is real

$$\varphi(k_1,k_2,z) = -\varphi^*(-k_1,-k_2,z)$$

so (17) reduces to

$$\left.\begin{aligned}
&\Psi(x,y,z,t) = \cos(nz-\omega t) - \frac{n^2L}{4\pi}\iiint (n^2-k^2)^{-\frac{1}{2}}\varphi(k_1,k_2,z')\,e^{+i(k_1x+k_2y)}\times \\
&\qquad\qquad \times\sin[nz'+(n^2-k^2)^{\frac{1}{2}}|z-z'|-\omega t]\,dk_1\,dk_2\,dz' \quad \text{for } n>k \\
&=\cos(nz-\omega t) + \frac{n^2L}{4\pi}\iiint (k^2-n^2)^{-\frac{1}{2}}\varphi(k_1,k_2,z')\,e^{+i(k_1x+k_2y)-(k^2-n^2)^{\frac{1}{2}}|z-z'|}\times \\
&\qquad\qquad \times\cos(nz'-\omega t)\,dk_1\,dk_2\,dz' \quad \text{for } n<k.
\end{aligned}\right\} \tag{18}$$

We shall now assume that the receiver responds only to incoming waves and that it is insensitive to any waves reflected from the medium behind it. This assumption is equivalent to putting $\varphi = 0$ when $z' > z$, so we can now remove the modulus signs surrounding $z-z'$ in (18).

VOL. 2 1952

The wave intensity is proportional to the time average over one cycle of $\Psi \int_0^t \nabla \Psi \, dt$, which can be calculated from (18). The resulting expression for $E(x, y, z)$ is, to the first order in φ,

$$\left.\begin{aligned} E(x,y,z) &= \frac{L}{4\pi} \iiint \frac{n\,[n + (n^2 - k^2)^{\frac{1}{2}}]}{(n^2 - k^2)^{\frac{1}{2}}} \varphi(k_1, k_2, z')\, e^{i(k_1 x + k_2 y)} \times \\ &\quad \times \sin\{[n - (n^2 - k^2)^{\frac{1}{2}}]\,[z - z']\}\, dk_1\, dk_2 \quad \text{for } n > k \\ \text{and} \quad &= \frac{L}{4\pi} \iiint \frac{n^2}{(k^2 - n^2)^{\frac{1}{2}}} \varphi(k_1, k_2, z')\, e^{i(k' x + k_2 y)}\, e^{-(k^2 - n^2)^{\frac{1}{2}}(z - z')} \times \\ &\quad \times \left\{\cos n(z - z') - \frac{(k^2 - n^2)^{\frac{1}{2}}}{n} \sin n(z - z')\right\} dk_1\, dk_2 dz' \quad \text{for } n < k. \end{aligned}\right\} \tag{19}$$

So from (16)

$$\left.\begin{aligned} I(k_1, k_2) &= \frac{1}{4} \iint \frac{n^2[n + (n^2 - k^2)^{\frac{1}{2}}]^2}{n^2 - k^2} \overline{\varphi(k_1, k_2, z')\, \varphi(-k_1, -k_2, z'')} \times \\ &\quad \times \sin\{[n - (n^2 - k^2)^{\frac{1}{2}}]\,[z_0 - z']\} \sin\{[n - (n^2 - k^2)^{\frac{1}{2}}]\,[z_0 - z'']\}\, dz'\, dz'' \quad \text{for } n > k \\ \text{and} \quad &= \frac{1}{4} \iint \frac{n^4}{(k^2 - n^2)} \overline{\varphi(k_1, k_2, z')\, \varphi(-k_1, -k_2, z'')}\, e^{-(k^2 - n^2)^{\frac{1}{2}}(2z_0 - z' - z'')} \times \\ &\quad \times \left\{\cos n(z_0 - z') - \frac{(k^2 - n^2)^{\frac{1}{2}}}{n} \sin n(z_0 - z')\right\} \times \\ &\quad \times \left\{\cos n(z_0 - z'') - \frac{(k^2 - n^2)^{\frac{1}{2}}}{n} \sin n(z_0 - z'')\right\} dz'\, dz'' \quad \text{for } n < k. \end{aligned}\right\} \tag{20}$$

Now let $\zeta = z' - z''$ and $Z = z' + z''$ so that $2dz\,dz' = d\zeta\, dZ$ and let $Q(k_1, k_2, \zeta, Z) = \overline{\varphi(k_1, k_2, z')\, \varphi(-k_1, -k_2, z'')}$. Q will then describe the statistical properties of the refractive index field. The dependence of Q on Z describes the variation of the statistical properties of the field in the z-direction. (We have assumed that the field is homogeneous in the x- and y-directions.) If the field is also homogeneous in the z-direction then Q will not depend on Z and will be related to the autocovariance $S(\xi\, \eta, \zeta)$ of the refractive index field by

$$Q(k_1, k_2, \zeta, Z) = \frac{1}{(2\pi)^2} \iint S(\xi, \eta, \zeta)\, e^{i(k_1 \xi + k_2 \eta)}\, d\xi\, d\eta. \tag{21}$$

With the introduction of ζ, Z and Q, (20) becomes

$$\left.\begin{aligned} I(k_1, k_2) &= \frac{1}{16} \iint \frac{n^2[n + (n^2 - k^2)^{\frac{1}{2}}]^2}{(n^2 - k^2)} Q(k_1, k_2, \zeta, Z) \times \\ &\quad \times \{\cos[n - (n^2 - k^2)^{\frac{1}{2}}]\,\zeta - \cos([n - (n^2 - k^2)^{\frac{1}{2}}]\,[2z_0 - Z])\}\, d\zeta\, dZ \quad \text{for } n > k \\ \text{and} \quad &= \frac{1}{16} \iint \frac{n^4}{(k^2 - n^2)} Q(k_1, k_2, \zeta, Z)\, e^{-(k^2 - n^2)^{\frac{1}{2}}(2z_0 - Z)} \times \\ &\quad \times \left\{\frac{k^2}{n^2} \cos n\zeta + \left(2 - \frac{k^2}{n^2}\right) \cos n(2z_0 - Z) - \frac{2(k^2 - n^2)^{\frac{1}{2}}}{n} \sin(2z_0 - Z)\right\} d\zeta\, dZ \\ &\qquad \text{for } n < k. \end{aligned}\right\} \tag{22}$$

This is the result of the diffraction theory. We shall discuss it when we have derived the result of the ray theory to compare with it.

VOL. 2 1952

The Ray Treatment

It is easy to calculate quantities such as the lateral spectrum for the fluctuation in intensity by the methods of geometrical optics. Such a calculation will be expected to agree with that of the previous section in the limit of infinite wave-number. The great simplicity of the ray treatment makes it attractive for practical problems where the wave-treatment may be made difficult by the complexity of the geometry or the form of the incident wave. It therefore seems desirable to compare the results of the ray- and wave-treatments in this case, so that the limitations of the ray treatment may be clearly understood.

We shall assume from the outset (and the conditions for the validity of this assumption can readily be calculated from the following work) that the lateral displacement of a ray is so small that the horizontal gradient of refractive index at a point on a ray can always be equated to that at the neighbouring point on the straight line through the receiver in the direction of the incoming beam.

As in the wave treatment we take the incident wave to approach along the z-axis. We now denote the x- and y-components of the unit tangent to the ray at (x, y, z) by $q_1(x, y, z)$ and $q_2(x, y, z)$. Since q_1 and q_2 are very small, the law of refraction can then be written

$$\left.\begin{aligned} \frac{\partial q_1}{\partial z} &= \frac{\partial \mu(x, y, z)}{\partial x} = -\frac{1}{2}\frac{\partial F(x, y, z)}{\partial x} \\ \frac{\partial q_2}{\partial z} &= \frac{\partial \mu(x, y, z)}{\partial y} = -\frac{1}{2}\frac{\partial F(x, y, z)}{\partial y} \end{aligned}\right\} \tag{23}$$

with the definition of $F(x, y, z)$ given by (1).

The x- and y-components of the lateral displacement of the ray, $s_1(x, y, z)$ and $s_2(x, y, z)$ will then be given by

$$\left.\begin{aligned} s_1(x, y, z) &= \int_0^z q_1(x, y, z')\, dz' = -\frac{1}{2}\int_0^z\int_0^{z'''} \frac{\partial F(x, y, z')}{\partial x}\, dz'\, dz''' \\ s_2(x, y, z) &= \int_0^z q_2(x, y, z')\, dz' = -\frac{1}{2}\int_0^z\int_0^{z'''} \frac{\partial F(x, y, z')}{\partial y}\, dz'\, dz'''. \end{aligned}\right\} \tag{24}$$

Now, according to ray theory, rays are streamlines for energy, so the intensity fluctuation at any point will be equal to minus the divergence of the lateral ray-displacement at that point; that is

$$E(x, y, z) = +\frac{1}{2}\int_0^z\int_0^{z'''} \left(\frac{\partial^2 F(x, y, z')}{\partial x^2} + \frac{\partial^2 F(x, y, z')}{\partial y^2}\right) dz'\, dz'''. \tag{25}$$

φ can now be introduced as in (8) so that (25) becomes

$$E(x, y, z_0) = \frac{L}{4\pi}\int_{\infty}^{\infty}\int_{\infty}^{\infty}\int_0^{z_0}\int_0^{z'''} k^2 \varphi(k_1, k_2, z')\, e^{-i(k_1 x + k_2 y)}\, dz'\, dz'''\, dk_1\, dk_2 \tag{26}$$

so by (16)

$$\left.\begin{aligned} I(k_1,k_2) &= \frac{1}{4}\int_0^{z}\int_0^{z}\int_0^{z^{IV}}\int_0^{z'''} k^4\,\overline{\varphi(k_1,k_2,z')\,\varphi(-k_1,-k_2,z'')}\,dz''\,dz'\,dz^{IV}\,dz''' \\ &= \frac{1}{4}\int_0^{z}\int_0^{z_0} k^4\,(z_0-z')\,(z_0-z'')\,\overline{\varphi(k_1,k_2,z')\,\varphi(-k_1,-k_2,z'')}\,dz''\,dz'. \end{aligned}\right\} \quad (27)$$

As in the diffraction case Q may be considered zero except in the region between $z=0$ and $z=z_0$. This enables us to take the range of integration in (27) to be infinite, so

$$I(k_1,k_2) = \frac{1}{32}\iint k^4\{(2z_0-Z)^2-\zeta^2\}\,Q(k_1,k_2,\zeta,Z)\,dZ\,d\zeta. \quad (28)$$

COMPARISON OF THE TWO THEORIES

Observe first that when k is vanishingly small compared with n,

$$(n^2-k^2)^{\frac{1}{2}} \sim n - k^2/2n$$

so

$$\cos[n-(n^2-k^2)^{\frac{1}{2}}]\,\zeta \sim 1-\frac{k^4\zeta^2}{8n^2}$$

and

$$\cos\{[n-(n^2-k^2)^{\frac{1}{2}}]\,[2z_0-Z]\} \sim 1-\frac{k^4(2z_0-Z)^2}{8n^2};$$

thus (22) comes to coincide with (28) as would be expected.

We shall now examine the behaviour of (22) when k is not so small. The dependence of Q on ζ will be such that it will yield a negligible contribution to the integral when ζ is greater than ζ_0 the scale of the variations in the refractive index. So, if, as will frequently be the case, components of the spectrum are being examined for which the wave number k is such that

$$\frac{k^2\zeta_0}{8n} \ll 1 \;i.e.\;\zeta_0 \gg \frac{k^2\zeta_0^2}{8n}, \quad (29)$$

we may replace $\cos[n-(n^2-k^2)^{\frac{1}{2}}]\zeta$ by 1 in (22). (Note that $k\zeta_0$ is the ratio of the scale of the turbulence to the size of the fluctuations which are being observed.) The expression for $I(k_1,k_2)$ when $n \gg k$ then becomes

$$I(k_1,k_2) = \frac{k^2}{4}\int_0^{2z_0}\int_\infty Q(k_1,k_2,\zeta,Z)\,d\zeta\left\{1-\cos\frac{k^2[2z_0-Z]}{2n}\right\}dZ. \quad (30)$$

The term $\left\{1-\cos\frac{k^2[2z_0-Z]}{2n}\right\}$ expresses the relative weight given to variation in refractive index at a distance $z_0-Z/2$ from the receiver.

If the waves are propagated a distance $z_0(\gg\zeta_0)$ through a homogeneous medium, *i.e.* one in which Q does not depend on Z,

$$I(k_1,k_2) = \frac{n^2}{4}\left\{2z_0-\frac{2n}{k^2}\sin\frac{k^2z_0}{n}\right\}\int Q(k_1,k_2,\xi)\,d\xi \quad (31)$$

which is proportional to z_0^3 for small $z_0\left(\frac{k^2z_0}{n}\ll 1\right)$ and to z_0 for large z_0 $\left(\text{when } \frac{k^2z_0}{n}\gg 1\right)$.

It is, perhaps, worth making one very particular substitution in (31) for the sake of illustration. Suppose

$$S(\xi, \eta, \zeta) = e^{-(\xi^2 + \eta^2 + \zeta^2)/l^2}$$

then

$$\int Q(k_1, k_2, \zeta)\, d\zeta = \frac{1}{(2\pi)^2} \iint e^{-(\xi^2 + \eta^2 + \zeta^2)/l^2}\, e^{i(k_1 \xi + k_2 \eta)}\, d\xi\, d\eta\, d\zeta$$

$$= \frac{l^2}{4\pi} e^{-k^2 l^2/4}$$

so

$$I(k_1, k_2) = \frac{n^2 l^2}{8\pi} \left\{ z_0 - \frac{n}{k^2} \sin \frac{k^2 z_0}{n} \right\}.$$

It will be seen from (31) that the ray theory is only adequate provided $\frac{k^2 z_0}{n} \ll 1$. It should be noted that contributions to $I(k_1, k_2)$ at a particular wave number arise entirely from that same wave number in the refractive index spectrum $\varphi(k_1, k_2, z)$. It is thus permissible to apply the ray theory to components of $I(k_1, k_2)$ which satisfy $\frac{k^2 z_0}{n} \ll 1$ even though variations of refractive index on a smaller scale are present.

Acknowledgement—The author wishes to thank F. URSELL for helpful discussions during the work.

VOL. 2 1952

REFERENCES

[1] DAHL, H. and DEVIK, O.; Nature, Lond. 1937 **139** 550–551. [2] BOVSHEVEROV, V. M. and KRASSILNIKOV, V. A.; C. R. Acad. Sci. URSS 1941 **32** 44–46. [3] KRASSILNIKOV, V. A.; C.R. Acad. Sci. URSS 1945 **46** 103–104. [4] KNUDSEN, V. O.; J. Acoust. Soc. Amer. 1946 **18** 90–96. [5] KRASSILNIKOV, V. A. and IVANHOV-SHITS, K. M.; Doklady Akad. Nauk. SSSR 1949 **67** 639–642. [6] RICHARDSON, E. G.; Proc. Roy. Soc. A. 1950 **203** 149–164. [7] LIEDENTOPF, H. and WISSHAK; Optik 1948 **3** 430–443. [8] MEGAW, E. C. S.; Nature, Lond. 1950 **166** 1100–1104. [9] BUTLER, H. E.; Nature, Lond. 1951 **167** 187. [10] KRASSILNIKOV, V. A.; Akad. Nauk. SSSR, Izvestiya Ser. Geogr. i. Geofiz. 1949 **13** 33–57. [11] MOTT, N. F. and MASSEY, H. S. W.; The Theory of Atomic Collisions (Second Edition), p. 116. Oxford, 1949.

Journal of Atmospheric and Terrestrial Physics, 1951, Vol. 2, pp. 22 to 31. Pergamon Press Ltd., London

Über die Entstehung von Oberwellen in der Ionosphäre

KARL FÖRSTERLING und HANS-OTTO WÜSTER

(Eingegangen am 10 April 1951)

ABSTRACT

On the generation of higher harmonics in the ionosphere: This paper continues previous studies on the propagation of plane electromagnetic waves in an inhomogeneous medium with plane statification. Near a zero of the dielectric constant, the electric field-component parallel to the plane of incidence shows a singularity, and it becomes infinite if the damping vanishes. While, for the vicinity of a simple zero, that fact can be proved only by developments into power-series, the case of a double zero allows to derive the singularities by a solution of MAXWELLS equations in terms of known functions. The singularity, thus stated, causes the generation of higher harmonics not radiated by the transmitter.

Bei der Untersuchung der Ausbreitung von ebenen elektromagnetischen Wellen in einem inhomogenen Medium mit ebener Schichtung lassen sich die MAXWELLschen Gleichungen durch geeignete Wahl der Koordinaten auf zwei gewöhnliche Differentialgleichungen zurückführen [1]. Wir lassen die Einfallsebene der Welle mit der xz-Ebene eines rechtwinkligen Koordinatensystems zusammenfallen und legen die z-Achse in die Normale zur Schicht. Dann sind die Feldstärken von der dritten Koordinate y unabhängig und die Dielektrizitätskonstante ε des inhomogenen Mediums ist eine Funktion der Koordinate z allein. Die Zeitabhängigkeit der Feldstärken wollen wir zunächst als periodisch von der Form $e^{i\omega t}$ ansehen. Um den lästigen Faktor ω/c zu unterdrücken, wollen wir alle Gleichungen in reduzierten Koordinaten x und z anschreiben, die mit den geometrischen Koordinaten x^* und z^* durch die Beziehungen

$$x = \frac{\omega}{c}\, x^*, \qquad z = \frac{\omega}{c}\, z^*$$

zusammenhängen.

Die MAXWELLschen Gleichungen in ihrer zeitfreien Form

$$i\,\varepsilon\,\mathfrak{E}_y = \frac{\partial \mathfrak{H}_x}{\partial z} - \frac{\partial \mathfrak{H}_z}{\partial x} \qquad -i\,\mathfrak{H}_y = \frac{\partial \mathfrak{E}_x}{\partial z} - \frac{\partial \mathfrak{E}_z}{\partial x}$$

$$i\,\mathfrak{H}_x = \frac{\partial \mathfrak{E}_y}{\partial z} \qquad i\,\varepsilon\,\mathfrak{E}_x = -\frac{\partial \mathfrak{H}_y}{\partial z}$$

$$i\,\mathfrak{H}_z = -\frac{\partial \mathfrak{E}_y}{\partial x} \qquad i\,\varepsilon\,\mathfrak{E}_z = \frac{\partial \mathfrak{H}_y}{\partial x}$$

lassen sich dann auf zwei voneinander unabhängige Differentialgleichungen für die zur Einfallsebene senkrechten Komponenten der Feldstärken zurückführen.

$$\Delta\,\mathfrak{E}_y + \varepsilon\,\mathfrak{E}_y = 0$$

$$\Delta\,\mathfrak{H}_y - \frac{1}{\varepsilon}\,\frac{\partial \varepsilon}{\partial z}\,\frac{\partial \mathfrak{H}_y}{\partial z} + \varepsilon\,\mathfrak{H}_y = 0 \qquad \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}\,.$$

Die übrigen Komponenten der Feldstärken lassen sich nach den MAXWELLschen Gleichungen einfach aus den Lösungen dieser Gleichungen bestimmen.

Die Gleichungen für $\mathfrak{E}_y$ und $\mathfrak{H}_y$ lassen sich leicht in x und z separieren. Mit dem Ansatz

$$\mathfrak{E}_y = u(z)\, e^{i q x}, \qquad \mathfrak{H}_y = v(z)\, e^{i q x}$$

erhalten wir für u und v die Gleichungen:

$$\frac{d^2 u}{dz^2} + (\varepsilon - q^2)\, u = 0, \qquad \frac{d^2 v}{dz^2} - \frac{1}{\varepsilon}\,\frac{\partial \varepsilon}{dz}\,\frac{dv}{\partial z} + (\varepsilon - q^2)\, v = 0.$$

Wir wollen uns besonders der Untersuchung von v zuwenden. Im allgemeinen kann ε als langsam veränderlich angesehen werden, nur in der Nähe einer Nullstelle von ε kann dies nicht mehr vorausgesetzt werden, da hier die Größen $\left(\frac{1}{\varepsilon}\,\frac{d\sqrt{\varepsilon}}{dz}\right)^2$ und $\frac{1}{\sqrt{\varepsilon^3}}\,\frac{d^2\sqrt{\varepsilon}}{dz^2}$ nicht mehr klein gegen ε sind. In der Umgebung einer Nullstelle nähert man ε am besten durch einen Potenzansatz der Form αz^k an, der als erstes Glied einer Taylorentwicklung um $z=0$ aufgefaßt werden kann. Ist der Bereich, in dem diese Näherung ausreicht, genügend groß, so können an seinen Rändern die asymptotischen Ausdrücke für die Lösungen der Differentialgleichungen eingesetzt werden, die mit den Lösungen für langsam veränderliche ε übereinstimmen.

Wir wollen die Gleichung speziell in der Nähe einer einfachen Nullstelle des reellen Teiles von ε untersuchen. Hier möge sich ε durch einen linearen Ansatz der Form

$$\varepsilon = \alpha z - i\beta = \alpha z'$$

genügend genau annähern lassen. Wie die Form dieses Ansatzes zeigt, legen wir den Nullpunkt unseres Koordinatensystems, dessen Wahl uns noch freigeblieben war, in die Nullstelle des reellen Teils von ε und berücksichtigen eine etwa vorhandene Dämpfung, indem wir z' als komplexe Größe mit konstantem negativem Imaginärteil auffassen. Die Konstanten α und β sind also positiv reelle Größen. Die Gleichung für v, die wir unter diesen Voraussetzungen erhalten

$$\frac{d^2 v}{dz'^2} - \frac{1}{z'}\,\frac{dv}{dz'} + (\alpha z' - q^2)\, v = 0$$

läßt sich in der Nähe der Stelle $z=0$ durch einen Potenzreihenansatz lösen [2].

Zwei unabhängige Lösungen der Gleichung sind gegeben durch:

$$v_1 = \sum_{h=2}^{\infty} a_h z'^h$$

$$v_2 = A\, v_1 \log z' + \sum_{h=0}^{\infty} b_h z'^h.$$

Für die Koeffizienten a_h and b_h gelten, wie sich leicht durch Einsetzen der Lösungen in die Differentialgleichung ergibt, die Rekursionsformeln:

$$a_{h+2} = -\frac{\alpha\, a_{h-1} - q^2 a_h}{h\,(h+2)}$$

$$b_{h+2} = -\frac{2 a_{h+2}\,(h+1) + \alpha\, b_{h-1} - q^2 b_h}{h\,(h+2)}.$$

Es ist nun zu untersuchen, wie sich das allgemeine Integral

$$v = A_1 v_1 + A_2 v_2$$

der Gleichung aus den beiden Grundlösungen aufbauen muß, um die gewünschte Lösung zu erhalten. Nehmen wir an, daß nur bei $z=+\infty$ Lichtquellen vorhanden sind, so muß für $z=-\infty$ die Lösung für v verschwinden, da ε hier negativ ist. Für $q=0$ können wir nun die Lösung unserer Gleichung in geschlossener Form hinschreiben:

$$v = z' Z_{\frac{2}{3}}\left(\frac{2}{3}\sqrt{\alpha z'^{\frac{3}{2}}}\right).$$

$Z_{\frac{2}{3}}$ bedeutet hier eine Zylinderfunktion der Ordnung 2/3. Da die Differentialgleichung stetig von dem Parameter q abhängt, gilt das gleiche auch von ihren Lösungen. Wir können also in unseren Potenzreihen zu $q=0$ übergehen und aus dem Verhalten der so gewonnenen Lösungen auf das Verhalten der Lösungen für $q \neq 0$ schließen. Für $q=0$ fällt der Logarithmus in v_2 fort und unsere Potenzreihen ergeben BESSELsche Funktionen:

$$v_1 \sim z' J_{\frac{2}{3}}\left(\frac{2}{3}\sqrt{\alpha z'^{\frac{3}{2}}}\right), \qquad v_2 \sim z' J_{-\frac{2}{3}}\left(\frac{2}{3}\sqrt{\alpha z'^{\frac{3}{2}}}\right).$$

Beide Lösungen genügen jedoch nicht der Grenzbedingung bei $z=-\infty$. Die richtige Lösung für $q=0$ ist vielmehr:

$$v \sim z' H^{(1)}_{\frac{2}{3}}\left(\frac{2}{3}\sqrt{\alpha z'^{\frac{3}{2}}} e^{2\pi i}\right) = z' \frac{i}{\sin\frac{2\pi}{3}}\left(e^{-i\frac{2\pi}{3}} J_{\frac{2}{3}}\left(\frac{2}{3}\sqrt{\alpha z'^{\frac{3}{2}}} e^{2\pi i}\right) - J_{-\frac{2}{3}}\left(\frac{2}{3}\sqrt{\alpha z'^{\frac{3}{2}}} e^{2\pi i}\right)\right).$$

Für $q=0$ setzt sich also die gewünschte Lösung linear aus *beiden* Grundlösungen zusammen. Aus den oben genannten Gründen muß sich dann aber auch im Falle $q \neq 0$ die gewünschte Lösung aus *beiden* Grundlösungen aufbauen. Es tritt also in der Entwicklung von $\mathfrak{H}_y$ das Glied $z'^2 \log z'$ auf.

Für große q läßt sich die Lösung der Gleichung für v angenähert durch bekannte Funktionen ausdrücken. Wir gehen zu diesem Zweck von v zu einer Größe $\chi(z)$ über, die wir durch

$$v(z) = \sqrt{\varepsilon(z)}\,\chi(z)$$

definieren wollen. Die Differentialgleichung für χ lautet dann:

$$\left(\varepsilon - q^2 - \sqrt{\varepsilon}\,\frac{\partial^2 \frac{1}{\sqrt{\varepsilon}}}{dz^2}\right)\chi + \frac{d^2\chi}{dz^2} = 0$$

und für den linearen Ansatz $\varepsilon = \alpha z'$ lautet sie:

$$\left(\alpha z' - q^2 - \frac{3}{4z'^2}\right)\chi + \frac{d^2\chi}{dz^2} = 0.$$

Für genügend große q kann man den Bereich von z in zwei Gebiete einteilen:

Für $|z| < \bar{z}$ sei $|\alpha z'|$ gegen q^2 zu vernachlässigen und für $|z| > \bar{z}$ sei $\left|\frac{3}{4z'^2}\right|$ gegen q^2 zu vernachlässigen. $\bar{z}$ ist dabei eine passend gewählte Konstante. Dann lautet für $|z| < \bar{z}$, also für den Bereich, in dem die uns interessierende Stelle $\varepsilon = 0$ liegt, die Differentialgleichung für χ

$$\frac{d^2\chi}{dz'^2} - \left(q^2 + \frac{3}{4z'^2}\right)\chi = 0.$$

Die Lösung muß bei $|z| = \bar{z}$ in die außerhalb dieses Bereiches geltende übergehen, die für $z = -\infty$ verschwindet

$$\chi^{(\nu)} = \sqrt{z'}\, H_1^{(\nu)}(i\,q\,z'), \qquad \nu = 1,2.$$

Entwickelt man nun die Lösungen in eine Potenzreihe, so ist ($|z| \ll 1$)

$$H_1^{(\nu)}(i\,q\,z') = J_1(i\,q\,z') \pm i\,N_1(i\,q\,z') \sim -\frac{q\,z'}{\pi}(1+\cdots)\log z' - \frac{2}{\pi\,q\,z'}(1+\cdots)$$

$$\chi \sim \frac{q\,z'^{\frac{3}{2}}}{\pi}(1+\cdots)\log z' + \frac{2}{\pi\,q\,z'^{\frac{1}{2}}}(1+\cdots)$$

$$\mathfrak{H}_y \sim \frac{q\,z'^2}{\pi}\log z' + \frac{2}{\pi q} + \cdots$$

Auch hier ist also in der Lösung für $\mathfrak{H}_y$ das Glied $z'^2 \log z'$ enthalten, und die Entwicklung von $\mathfrak{H}_y$ enthält ein konstantes Glied.

Wegen des durch die MAXWELLschen Gleichungen gegebenen Zusammenhanges zwischen $\mathfrak{H}_y$ und den Komponenten $\mathfrak{E}_x$ und $\mathfrak{E}_z$ enthält dann aber die Entwicklung von $\mathfrak{E}_x$ das Glied $\log z'$ und die Entwicklung von $\mathfrak{E}_z$ das Glied $1/z'$. Bei Wegfall der Dämpfung wird nun z' reell und damit werden $\mathfrak{E}_x$ und $\mathfrak{E}_z$ für $z = 0$ unendlich. Für den Fall einer mehrfachen Nullstelle verstärkt sich dieses Verhalten noch: An einer k-fachen Nullstelle von ε hat $\mathfrak{E}_z$ einen k-fachen Pol und $\mathfrak{E}_x$ einen $(k-1)$-fachen Pol. Die Ladungsdichte berechnet sich nach der Gleichung

$$\operatorname{div} \mathfrak{E} = 4\pi\varrho.$$

Sie würde also an der Stelle $\varepsilon = 0$ beim linearen Ansatz unendlich wie $1/z^2$. An einer k-fachen Nullstelle von ε hätte sie sogar einen $(k+1)$-fachen Pol. Ein solches Verhalten ist aber unmöglich.

Um zu einem physikalisch richtigen Ergebnis zu kommen, haben wir die Berechnung von ε einer genaueren Betrachtung unterzogen. Da in der Ionosphäre gebundene Elektronen keine Rolle spielen, definieren wir hier den Vektor der elektrischen Verschiebung $\mathfrak{D}$ durch:

$$\frac{\partial \mathfrak{D}}{\partial t} = \frac{\partial \mathfrak{E}}{\partial t} + 4\pi\,\mathfrak{i}.$$

ε erhalten wir dann aus der Gleichung:

$$\mathfrak{D} = \varepsilon\,\mathfrak{E},$$

während sich die Stromdichte $\mathfrak{i}$ nach der LORENTZschen Elektronentheorie aus der Anzahl N der Elektronen und ihrer Geschwindigkeit $\mathfrak{v}$ berechnet.

$$\mathfrak{i} = N e\,\mathfrak{v}.$$

e ist dabei die Ladung des Elektrons.

Die Geschwindigkeit $\mathfrak{v}$ wiederum berechnet sich nach den Bewegungsgleichungen der Mechanik:

$$m\frac{d\mathfrak{v}}{dt} = e\,\mathfrak{E}.$$

Die rechte Seite dieser Gleichung stellt die vom Felde auf das Elektron ausgeübte Kraft dar. Wir nehmen an, daß unsere oben abgeleitete Lösung für $\mathfrak{H}_y$ insoweit

VOL. 2 1952

richtig ist, daß wir ein starkes und schnelles Anwachsen von $\mathfrak{E}_x$ und $\mathfrak{E}_z$ in der Nähe von $z=0$ voraussetzen können. Das bedeutet dann aber, daß das Elektron Schwingungen sehr großer Elongation ausführt und auf seinem Wege während einer Schwingung Gebiete mit sehr verschiedenen Feldstärkeamplituden berührt. Wir haben daher, im Gegensatz zu der sonst üblichen Methode, wohl zu unterscheiden zwischen der Feldstärke $\mathfrak{E}$, die sich auf einen festen Punkt des Raumes bezieht und der Feldstärke am Orte des Elektrons, die wir mit $[\mathfrak{E}]$ bezeichnen wollen. (Von dem Umstande, daß $\mathfrak{E}$ nach der LORENTZschen Elektronentheorie strenggenommen einen Mittelwert über ein physikalisch unendlich kleines Volumenelement bedeutet, wollen wir hier absehen.) Während man sonst diesen Unterschied vernachlässigen kann und ohne weiteres in die Bewegungsgleichung $\mathfrak{E}$ aus den MAXWELLschen Gleichungen einsetzen kann, müssen wir in unserem Falle die wirkliche Feldstärke am Orte des Elektrons in die Bewegungsgleichungen einsetzen, die der Beziehung

$$[\mathfrak{E}] = \mathfrak{E}\left(x + \int_0^t v_x\, dt,\ z + \int_0^t v_z\, dt\right)$$

genügt. Nun ist die direkte Berechnung dieser Größe nicht möglich, da wir dazu erst die Geschwindigkeit $\mathfrak{v}$ kennen müßten; $\mathfrak{v}$ erhalten wir jedoch erst nach Kenntnis von $[\mathfrak{E}]$ aus den Bewegungsgleichungen. Wir sind also auf ein Näherungsverfahren angewiesen. Da in der Nähe von $z=0$ $|\mathfrak{E}_z| \gg |\mathfrak{E}_x|$ ist dort auch $|v_z| \gg |v_x|$, wir lassen aus diesem Grunde im Argument von $[\mathfrak{E}]$ das Glied $\int_0^t v_x\, dt$ fort.

Obwohl wir in der F-Schicht der Ionosphäre eine stärkere Dämpfung durch Stöße der Elektronen mit den Atomen der Schicht nicht voraussetzen können und daher eine solche auch in den Bewegungsgleichungen vernachlässigt haben, nehmen wir an, daß eine Dämpfung vorhanden sei, über deren Größe wir zunächst noch nichts aussagen können. Es wird sich zeigen, daß diese Dämpfung dadurch zustande kommt, daß in $[\mathfrak{E}]$ außer der in der einfallenden Welle auftretenden Grundschwingung auch deren harmonische Oberschwingungen auftreten. Die durch diese Oberschwingungen der Grundschwingung entzogene Energie muß sich dann in einer zusätzlichen Dämpfung der Grundschwingung bemerkbar machen. Wir wollen hier nur die z-Komponente von $[\mathfrak{E}]$ unter Zugrundelegung des linearen Ansatzes für ε berechnen, die sich nach dem oben Gesagten folgendermaßen schreiben läßt:

$$[\mathfrak{E}_z] = e^{i q x + i \omega t}\, w\left(z + \int_0^t v_z\, dt\right)$$

$$w(z) = \frac{A}{z'} + \varphi(z').$$

A ist eine Konstante und φ eine reguläre Funktion von z'. Der physikalisch wirksame Bestandteil dieser Größe wird durch den Realteil dargestellt. Für diesen erhalten wir:

$$\mathfrak{E}_z = \eta_1' \cos \omega t + \eta_1'' \sin \omega t$$

mit

$$\eta_1' = \frac{1}{z^2 + \frac{\beta^2}{\alpha^2}}\left(z\,\chi_1' - \frac{\beta}{\alpha}\chi_1''\right) + \chi_2', \qquad \eta_1'' = -\frac{1}{z^2 + \frac{\beta^2}{\alpha^2}}\left(z\,\chi_1'' + \frac{\beta}{\alpha}\chi_1'\right) - \chi_2''$$

und den Abkürzungen

$$A\,e^{i\,q\,x} = \chi_1' + i\,\chi_1'', \qquad \varphi(z')\,e^{i\,q\,x} = \chi_2' + i\,\chi_2''.$$

Hiermit gehen wir zunächst in die Bewegungsgleichungen ein, um eine erste Näherung v_z zu erhalten

$$\frac{dv_z}{dt} = \frac{e}{m}\,(\eta_1' \cos\omega t + \eta_1'' \sin\omega t).$$

Die Lösung für v_z lautet:

$$v_z = v_1' \cos\omega t + v_1'' \sin\omega t$$

mit den Abkürzungen

$$v_1' = -\frac{e}{m\,\omega}\,\eta_1'' \qquad v_1'' = \frac{e}{m\,\omega}\,\eta_1'.$$

Wie man leicht sieht, würden v_1' und v_1'' in dieser Näherung bei Wegfall der Dämpfung unendlich werden.

Um nun $[\mathfrak{E}_z]$ zu bilden, haben wir in $\mathfrak{E}_z$ überall an Stelle von z die Größe $z + \int_0^t v_z\,dt$ einzusetzen. Dieser Unterschied macht sich vor allen Dingen in dem Glied A/z' in $\mathfrak{E}_z$ bemerkbar, wir wollen ihn in der regulären Funktion φ vernachlässigen. Wir erhalten dann für $[\mathfrak{E}_z]$:

$$[\mathfrak{E}_z] = \frac{1}{\left(z + \int_0^t v_z\,dt\right)^2 + \frac{\beta^2}{\alpha^2}}\left(\left(z + \int_0^t v_z\,dt\right)\chi_1' - \frac{\beta}{\alpha}\,\chi_1''\right)\cos\omega t + \chi_2' \cos\omega t -$$

$$- \frac{1}{\left(z + \int_0^t v_z\,dt\right)^2 + \frac{\beta^2}{\alpha^2}}\left(\left(z + \int_0^t v_z\,dt\right)\chi_1'' + \frac{\beta}{\alpha}\,\chi_1'\right)\sin\omega t - \chi_2'' \sin\omega t.$$

Für $\int_0^t v_z\,dt$ haben wir nun einzusetzen:

$$\int_0^t v_z\,dt = \frac{v_1'}{\omega}\sin\omega t + \frac{v_1''}{\omega}\,(1 - \cos\omega t).$$

Damit wird aber:

$$[\mathfrak{E}_z] = \frac{\left(z + \frac{v_1'}{\omega}\sin\omega t + \frac{v_1''}{\omega}\,(1-\cos\omega t)\right)\chi_1' - \frac{\beta}{\alpha}\,\chi_1''}{\left(z + \frac{v_1'}{\omega}\sin\omega t + \frac{v_1''}{\omega}\,(1-\cos wt)\right)^2 + \frac{\beta^2}{\alpha^2}}\cos\omega t + \chi_2' \cos\omega t$$

$$\frac{\left(z + \frac{v_1'}{\omega}\sin\omega t + \frac{v_1''}{\omega}\,(1-\cos\omega t)\right)\chi_1' + \frac{\beta}{\alpha}\,\chi_1'}{\left(z + \frac{v_1'}{\omega}\sin\omega t + \frac{v_1'}{\omega}\,(1-\cos\omega t)\right)^2 + \frac{\beta^2}{\alpha^2}}\sin\omega t - \chi_2'' \sin\omega t.$$

Wir sehen aus der Form dieses Ausdrucks für $[\mathfrak{E}_z]$, daß $[\mathfrak{E}_z]$ zwar auch periodisch von der Periode $\tau = 2\pi/\omega$ ist, aber außer der Grundschwingung auch deren harmonische Oberschwingungen enthalten muß. Wir wollen zur Ermittlung dieser Oberschwingungen $[\mathfrak{E}_z]$ in eine Fourierreihe entwickeln:

$$[\mathfrak{E}_z] = [\eta_0'] + [\eta_1'] \cos\omega t + [\eta_2'] \cos 2\,\omega t + \cdots$$
$$+ [\eta_1''] \sin\omega t + [\eta_2''] \sin 2\,\omega t + \cdots.$$

VOL. 2 1952

Die Berechnung dieser Reihe gestaltet sich einfacher, wenn wir von der Entwicklung der Größe $[\mathfrak{E}_z] - \mathfrak{E}_z$ ausgehen:

$$[\mathfrak{E}_z] - \mathfrak{E}_z = [\eta_0'] + ([\eta_1'] - \eta_1') \cos \omega t + [\eta_2'] \cos 2\omega t + \cdots + ([\eta_1''] - \eta_1'') \sin \omega t + [\eta_2''] \sin 2\omega t + \cdots .$$

Die Koeffizienten der beiden Entwicklungen stimmen bis auf die Koeffizienten der Grundschwingung überein, da ja $\mathfrak{E}_z$ nur die Grundschwingung enthält. Wir erhalten für $[\mathfrak{E}_z] - \mathfrak{E}_z$:

$$[\mathfrak{E}_z] - \mathfrak{E}_z = \frac{\frac{v_1'}{\omega} \sin \omega t + \frac{v_1''}{\omega}(1 - \cos \omega t)}{\left(z^2 + \frac{\beta^2}{\alpha^2}\right)\left(\left(z + \frac{v_1'}{\omega} \sin \omega t + \frac{v_1''}{\omega}(1 - \cos \omega t)\right)^2 + \frac{\beta^2}{\alpha^2}\right)} \times$$
$$\times \left\{\left(\frac{\beta^2}{\alpha^2} - z^2 - z \frac{v_1'}{\omega} \sin \omega t - z \frac{v_1''}{\omega}(1 - \cos \omega t)\right)(\chi_1' \cos \omega t - \chi_1'' \sin \omega t)\right.$$
$$\left. + \left(2 z \frac{\beta}{\alpha} + \frac{\beta}{\alpha} \frac{v_1'}{\omega} \sin \omega t + \frac{\beta}{\alpha} \frac{v_1''}{\omega}(1 - \cos \omega t)\right)(\chi_1'' \cos \omega t + \chi_1' \sin \omega t)\right\}.$$

Die Koeffizienten der Grundschwingung in der Entwicklung dieser Größe sind:

$$[\eta_1'] - \eta_1' = \frac{1}{\pi} \int_{-\pi}^{+\pi} ([\mathfrak{E}_z] - \mathfrak{E}_z) \cos \omega t \, d(\omega t), \qquad [\eta_1''] - \eta_1'' = \frac{1}{\pi} \int_{-\pi}^{+\pi} ([\mathfrak{E}_z] - \mathfrak{E}_z) \sin \omega t \, d(\omega t).$$

Die Auswertung dieser Integrale ist sehr umständlich, wir geben daher nur das Ergebnis an:

$$[\eta_1'] - \eta_1' = \frac{\frac{\beta}{\alpha} \chi_1'' - z \chi_1'}{z^2 + \frac{\beta^2}{\alpha^2}}, \qquad [\eta_1''] - \eta_1'' = \frac{\frac{\beta}{\alpha} \chi_1' + z \chi_1''}{z^2 + \frac{\beta^2}{\alpha^2}}.$$

Mit diesem Ergebnis erhalten wir für die Koeffizienten der Grundschwingung in der Entwicklung von $[\mathfrak{E}_z]$:

$$[\eta_1'] = \chi_2', \qquad [\eta_1''] = -\chi_2''.$$

Hiermit wollen wir vor der Berechnung der Koeffizienten der Oberschwingungen in die Bewegungsgleichungen eingehen, um so eine bessere Näherung für v_z zu erhalten. Da in $[\mathfrak{E}_z]$ gerade der für $z = 0$ unendlich werdende Teil von $\mathfrak{E}_z$ wegfällt, bleibt in dieser Näherung auch v_z endlich.

Wir wollen auch in dieser Näherung schreiben:

$$v_z = v_1' \cos \omega t + v_1'' \sin \omega t.$$

Nur haben hier die Abkürzungen v_1' und v_1'' die Bedeutung:

$$v_1' = \frac{e}{m \omega} \chi_2'', \qquad v_1'' = \frac{e}{m \omega} \chi_2'.$$

Auch diese Größen bleiben also für $z = 0$ endlich. Mit diesen Werten für v_z erhalten wir:

$$\int_0^t v_z \, dt = \frac{e}{m \omega^2} \chi_2'' \sin \omega t + \frac{e}{m \omega^2} \chi_2' (1 - \cos \omega t).$$

Für die Koeffizienten der Oberschwingung erhalten wir dann:

$$[\eta_2'] = 2\frac{m}{e}\omega^2\frac{\chi_1'\chi_2' - \chi_1''\chi_2''}{\chi_2'^2 + \chi_2''^2}, \qquad [\eta_2''] = 2\frac{m}{e}\omega^2\frac{\chi_1'\chi_2'' + \chi_1''\chi_2'}{\chi_2'^2 + \chi_2''^2}.$$

Das konstante Glied der Fourierentwicklung verschwindet.

Die in $[\mathfrak{E}_z]$ auftretenden Oberschwingungen gehen nun durch die Bewegungsgleichungen auch auf die Geschwindigkeitskomponente v_z über:

$$v_z = \frac{e}{m\omega}\chi_2''\cos\omega t - \omega\frac{\chi_1'\chi_2'' + \chi_1''\chi_2'}{\chi_2'^2 + \chi_2''^2}\cos 2\omega t + \cdots +$$
$$+ \frac{e}{m\omega}\chi_2'\sin\omega t + \omega\frac{\chi_1'\chi_2' - \chi_1''\chi_2''}{\chi_2'^2 + \chi_2''^2}\sin 2\omega t + \cdots.$$

Bilden wir nun die Stromdichte i, so treten auch in dieser Größe und damit auch in $\mathfrak{D}$ die Oberschwingungen auf:

$$\mathfrak{i}_z = N e v_z = \mathfrak{j}_1'\cos\omega t + \mathfrak{j}_2'\cos 2\omega t + \cdots +$$
$$+ \mathfrak{j}_1''\sin\omega t + \mathfrak{j}_2''\sin 2\omega t + \cdots.$$

Die in der Stromdichte auftretenden Oberschwingungen induzieren Oberschwingungen in den Feldstärken, die nicht vom Sender ausgestrahlt werden, sondern erst in der Ionosphäre entstehen und dort eine entsprechende Strahlung hervorrufen. Wir bemühen uns zur Zeit, die Form der so entstehenden Oberschwingungen festzustellen und die Ergebnisse der gleichen Untersuchung bei der Ausbreitung unter Berücksichtigung des Magnetfeldes der Erde in eine Form zu bringen, die ihre Überprüfung im Experiment möglich machen. Der Einfluß der Singularität ist wahrscheinlich auf ihre nächste Umgebung beschränkt, so daß die Oberwellen nur unter günstigen Umständen beobachtet werden können.

Wenn die benutzte Wellenlänge gerade die Grenzwellenlänge bei senkrechter Inzidenz ist, also $\varepsilon = \alpha z^2$ gesetzt werden kann, so verhält sich $\mathfrak{E}_z$ wie $1/z^2$, so daß die Verhältnisse für die Entstehung von Oberwellen dann günstiger werden.

Zusatz von H. O. Wüster. Über das Verhalten der Feldstärken in der Nähe einer zweifachen Nullstelle von ε (Grenzwelle).

Bei dem im vorstehenden behandelten Fall $\varepsilon = \alpha z - i\beta$ läßt sich das Verhalten der Feldstärken bei $z = 0$ nur mittels komplizierter Betrachtungen übersehen. Bei dem überraschenden Resultat ist es daher wichtig, daß für den Fall, daß ε sich wie αz^2 bei beliebigem q verhält, die Lösung sich in geschlossener Form angeben läßt; das Verhalten bei $z = 0$ läßt sich dann durch bekannte Reihenentwicklungen ableiten.

In der Nähe einer zweifachen Nullstelle von ε, wie eine solche z.B. für die Grenzwelle in der Ionosphäre auftritt, ist es zweckmäßig, die Dielektrizitätskonstante ε durch einen Ansatz der Form

$$\varepsilon = \alpha z^2$$

darzustellen. Die Differentialgleichung für die aus $\mathfrak{H}_y$ abgeleitete Größe $v(z)$ lautet mit diesem Ausdruck für ε:

$$\frac{d^2v}{dz^2} - \frac{2}{z}\frac{dv}{dz} + (\alpha z^2 - q^2)v = 0.$$

VOL. 2 1952

Die Lösung dieser Gleichung kann durch bekannte Funktionen in geschlossener Form ausgedrückt werden [3]. Ein Grundsystem von Lösungen wird gegeben durch:

$$v_1 = \sqrt{z}\, W_{\frac{i q^2}{4\sqrt{\alpha}}, \frac{3}{4}}(i\sqrt{\alpha}\, z^2), \qquad v_2 = \sqrt{z}\, W_{-\frac{i q^2}{4\sqrt{\alpha}}, \frac{3}{4}}(-i\sqrt{\alpha}\, z^2).$$

Die Lösungen lassen sich also durch die der konfluenten hypergeometrischen Reihe nahe verwandten WHITTAKERschen Funktionen W ausdrücken. Die asymptotische Entwicklung der WHITTAKERschen Funktionen lautet [4]:

$$W_{\varkappa,\mu}(z) \approx e^{-\frac{z}{2}} z^{\varkappa}\left(1 + \sum_{n=1}^{N} \frac{[\mu^2-(\varkappa-\frac{1}{2})^2]\,[\mu^2-(\varkappa-\frac{3}{2})^2]\ldots[\mu^2-(\varkappa-n+\frac{1}{2})^2]}{n!\, z^n}\right) |\arg z| < \pi.$$

Die Lösung unserer Gleichung, die den Grenzbedingungen genügt, die sich ergeben, wenn man annimmt, daß nur bei $z = +\infty$ Lichtquellen vorhanden sind, lautet also:

$$v = A\sqrt{z}\, W_{\frac{i q^2}{4\sqrt{\alpha}}, \frac{3}{4}}(i\sqrt{\alpha}\, z^2\, e^{2\pi i}).$$

Der Übergang von $i\sqrt{\alpha}\, z^2$ zu $i\sqrt{\alpha}\, z^2 e^{2\pi i}$ ist notwendig, damit wir im Gültigkeitsbereich der asymptotischen Darstellung bleiben.

Läuft nämlich z von $+\infty$ über 0 nach $-\infty$, so läuft $\arg z$, das wegen der Zeitabhängigkeit in der Form $e^{i\omega t}$ immer kleiner als 0 sein muß, von 0 bis -2π und $\arg(i\sqrt{\alpha}\, z^2)$ von $+\pi/2$ über $-\pi/2$ nach $-3\pi/2$. Beim Übergang von negativen zu positiven z in unserer Lösung verlassen wir wieder den Gültigkeitsbereich der asymptotischen Darstellung und müssen mit Hilfe der Umlaufrelation für die WHITTAKERschen Funktionen zu der hier brauchbaren Lösung übergehen. Es ist

$$W_{\varkappa,\mu}(z\, e^{2\pi i}) = -(2\cos 2\pi\mu + e^{2i\pi\varkappa})\, W_{\varkappa,\mu}(z) + \frac{2i\pi\, e^{-i\pi\varkappa}}{\Gamma(\frac{1}{2}-\mu-\varkappa)\,\Gamma(\frac{1}{2}+\mu-\varkappa)} W_{-\varkappa,\mu}(z\, e^{-i\pi}),$$

die Umlaufrelation für die WHITTAKERsche Funktion.

Damit wird aber dann unsere Lösung für positive z:

$$v = A\sqrt{z}\left\{-e^{\frac{\pi q^2}{2\sqrt{\alpha}}} W_{\frac{i q^2}{4\sqrt{\alpha}}, \frac{3}{4}}(i\sqrt{\alpha}\, z^2) + \frac{2i\pi\, e^{\frac{\pi q^2}{4\sqrt{\alpha}}}}{\Gamma\left(\frac{5}{4} - \frac{i q^2}{4\sqrt{\alpha}}\right)\Gamma\left(-\frac{1}{4} - \frac{i q^2}{4\sqrt{\alpha}}\right)} W_{-\frac{i q^2}{4\sqrt{\alpha}}, \frac{3}{4}}(-i\sqrt{\alpha}\, z^2).\right.$$

Die reflektierte Welle wird durch

$$W_{\frac{i q^2}{4\sqrt{\alpha}}, \frac{3}{4}}(i\sqrt{\alpha}\, z^2)$$

dargestellt, während

$$W_{-\frac{i q^2}{4\sqrt{\alpha}}, \frac{3}{4}}(-i\sqrt{\alpha}\, z^2)$$

der einfallenden Welle entspricht. Für sehr kleine z wollen wir $\mathfrak{H}_y$ und die daraus abgeleiteten Größen

$$\mathfrak{E}_x = \frac{i}{\varepsilon}\frac{\partial \mathfrak{H}_y}{\partial z} \quad \text{und} \quad \mathfrak{E}_z = -\frac{i}{\varepsilon}\frac{\partial \mathfrak{H}_y}{\partial x}$$

in Potenzreihen entwickeln, wobei wir die Beziehungen [5]

$$W_{\varkappa,\mu}(z) = \frac{\Gamma(-2\mu)}{\Gamma(\frac{1}{2}-\mu-\varkappa)} z^{\mu+\frac{1}{2}} e^{-\frac{z}{2}} {}_1F_1\left(\mu-\varkappa+\frac{1}{2}; 2\mu+1; z\right) +$$
$$+ \frac{\Gamma(2\mu)}{\Gamma(\frac{1}{2}+\mu-\varkappa)} z^{-\mu+\frac{1}{2}} e^{-\frac{z}{2}} {}_1F_1\left(-\mu-\varkappa+\frac{1}{2}; -2\mu+1; z\right)$$

und

$${}_1F_1(a; c; z) = 1 + \frac{a}{c}\frac{z}{1!} + \frac{a(a+1)}{c(c+1)}\frac{z^2}{2!}$$

benutzen. Hiernach ist aber:

$$\mathfrak{H}_y = A\, e^{i q x + i\omega t}\left\{\frac{\sqrt{\pi}\, e^{-i\frac{5\pi}{8}}}{2\alpha^{\frac{1}{4}}\Gamma\left(\frac{5}{4}-\frac{i q^2}{4\sqrt{\alpha}}\right)}\left\{1 - z^2\frac{q^2}{2} - \alpha z^4\left(\frac{1}{4}+\frac{q^4}{8\alpha}\right)\right\} - \frac{4\sqrt{\pi}\, e^{i\frac{\pi}{8}}\alpha^{\frac{5}{8}}}{3\Gamma\left(\frac{5}{4}-\frac{i q^2}{4\sqrt{\alpha}}\right)} z^3 + \cdots\right\}$$

$$\mathfrak{E}_z = A\, q\, e^{i q x + i\omega t}\left\{\frac{\sqrt{\pi}\, e^{-i\frac{5\pi}{8}}}{2\alpha^{\frac{3}{4}}\Gamma\left(-\frac{1}{4}-\frac{i q^2}{4\sqrt{\alpha}}\right)}\left\{\frac{1}{z^2} - \frac{q^2}{2} - \alpha z^2\left(\frac{1}{4}+\frac{q^4}{8\alpha}\right)\right\} - \frac{4\sqrt{\pi}\, e^{i\frac{\pi}{8}}}{3\alpha^{\frac{3}{4}}\Gamma\left(\frac{5}{4}-\frac{i q^2}{4\sqrt{\alpha}}\right)} z + \cdots\right.$$

$$\mathfrak{E}_x = A\, e^{i q x + i\omega t}\left\{\frac{-\sqrt{\pi}\, e^{-i\frac{\pi}{8}}}{2\alpha^{\frac{3}{4}}\Gamma\left(-\frac{1}{4}-\frac{i q^2}{4\sqrt{\alpha}}\right)}\left\{\frac{q^2}{z} + \alpha z\left(1+\frac{q^4}{2\alpha}\right)\right\} - \frac{4\sqrt{\pi}\, e^{i\frac{5\pi}{8}}}{\alpha^{\frac{3}{4}}\Gamma\left(\frac{5}{4}-\frac{i q^2}{4\sqrt{\alpha}}\right)}\right\}.$$

VOL. 2 1952

An diesen Potenzreihen ist zu erkennen, warum die auch hier auftretenden Singularitäten für $q=0$, also senkrechte Inzidenz, verschwinden.

Wir danken der Notgemeinschaft der Deutschen Wissenschaft auch an dieser Stelle für die Ermöglichung der Durchführung der vorliegenden Arbeit.

Zusammenfassung

In der vorliegenden Arbeit wird in Anlehnung an frühere Arbeiten gezeigt, daß die elektrische Feldstärke parallel der Einfallsebene an einer Nullstelle der Dielektrizitätskonstanten ein singuläres Verhalten zeigt und bei verschwindender Dämpfung unendlich wird. Während diese Tatsache für die Umgebung einer einfachen Nullstelle nur durch Potenzreihenentwicklungen nachgewiesen werden kann, ist es im Falle einer doppelten Nullstelle möglich, die Singularitäten aus einer Darstellung der Lösung der Maxwellschen Gleichungen durch bekannte Funktionen zu gewinnen. Die so festgestellte Singularität gibt Anlaß zum Auftreten von Oberwellen, die nicht vom Sender ausgestrahlt werden.

Literatur

[1] Vgl. Försterlin, K.; Arch. d. el. Übertr. 1949 **3** 115–120. [2] Vgl. Försterlin, K.; *loc.cit.* S. 119. [3] Vgl. z. B. Kamke, E.; Differentialgleichungen (2. Aufl.) S. 475. [4] Vgl. z. B. Whittaker, E. T. and Watson, G. N.; Modern Analysis (4th ed.) S. 342. [5] Vgl. z. B. Whittaker, E. T. and Watson, G. N.; *loc.cit.* Kap. 16.

Journal of Atmospheric and Terrestrial Physics, 1951, Vol. 2, pp. 32 to 37. Pergamon Press Ltd., London

An apparatus for simultaneous registration of potential gradient and air-earth current

(Description and first results)

H.-W. KASEMIR

Luftelektrische Forschungsstelle Buchau a. F. des Observatoriums Friedrichshafen, Landeswetterdienst Württemberg-Hohenzollern

(Received 1 November 1950)

ABSTRACT

An apparatus for the simultaneous continuous recording of the atmospheric electric potential gradient and the air-earth current is described. The CR constants are chosen in such a way that minor changes of the atmospheric electric elements in question are suppressed and the diurnal variations can readily be seen from the records. Modern technical means such as six-channel recorders, electrometer tubes, and low loss cables make it possible to register a great range of potential gradient and air-earth current. Moreover the apparatus is foolproof and convenient as to dimensions and weight so that it can be used on mountain observatories and expeditions without difficulty. Preliminary results obtained on mountain peaks are described, which will help in the interpretation of such records.

I. INTRODUCTION

During recent investigations on atmospheric electricity it has been established that, when recording only the potential gradient, it is very difficult to obtain proper knowledge on the atmospheric electric state of the atmosphere [1]. Recording of further elements, *e.g.* air-earth current or conductivity, involves many difficulties and only a small number of observers has been able to overcome them. Besides, when a second element was observed, usually the conductivity was recorded and it was only at Kew where the air-earth current was registered for a longer period. It is possible to calculate the air-earth current i from the potential gradient E and the conductivity λ from the equation $i = \lambda \times E$, as long as OHMs Law holds. It is possible that this condition is not fulfilled during windy weather periods, *i.e.* when strong convectional currents prevail.

Moreover, the air-earth current seems to be the primary and therefore the most informative of the three atmospheric electric elements, since during periods of the stationary state it depends very little on the variations of the local conductivity, but mainly on the variations of the atmospheric electric generator; that is, the world wide thunderstorm activity. Therefore, the air-earth current is the most important of the atmospheric electric elements and it is worthwhile to try its continuous registration.

Following the ideas outlined above H. ISRAËL asked the author to develop an instrument for the simultaneous registration of potential gradient and air-earth current. Modern technical means make it possible to construct an apparatus that fulfills the conditions set by atmospheric electric investigators and is simple as to construction and operation. Also, its weight is so low that it invites scientists to take it along for work on expeditions and high mountain peaks. In the following

the instrument and first tests of its performance are described. Figure 1 shows the apparatus together with a six-channel recorder and Figure 2 shows the circuit diagram.

II. Registration of Potential Gradient

The potential gradient is picked up by a radioactive collector which does not work directly on the recorder as it does in the case of the electrostatic Benndorf-recorder but it controls the grid of an electrometer tube the plate current of which is registered by means of a two- or six-channel recorder. The scale of the recorder comprises a range of 0·1 mA. Its rugged construction makes the instrument convenient for use on expeditions where rough treatment cannot be avoided. Besides, up to six elements can be recorded on one strip of paper simultaneously so that it is possible to record not only potential gradient and air-earth current but in addition also meteorological elements such as temperature, humidity, sunshine, wind etc., a fact that facilitates studying of correlations.

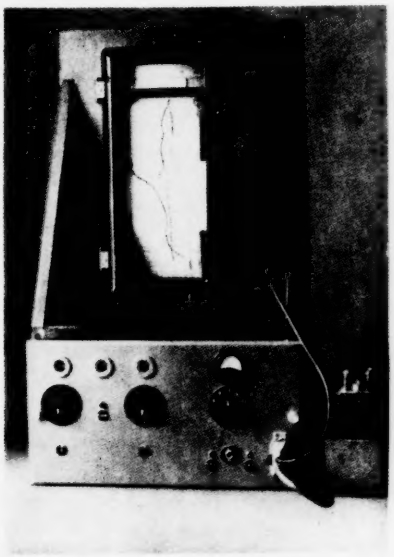


Fig. 1. Amplifier and six-channel recorder for the registration of air-earth current and potential gradient.

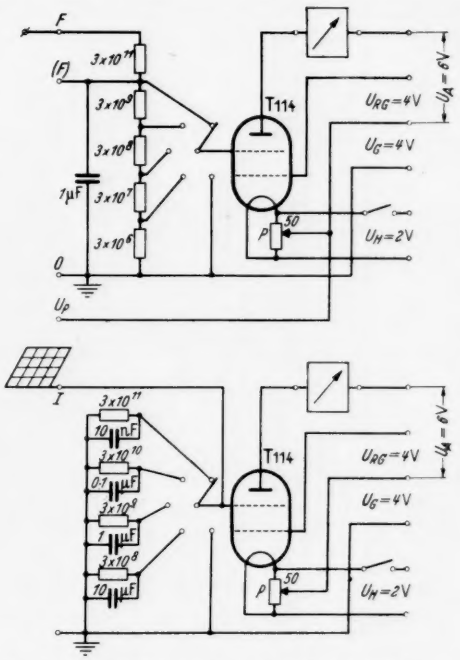


Fig. 2. Circuit diagram, above potential gradient, below air-earth current.

It can be inferred from the circuit diagram (Figure 2) that the electrometer tube works as a simple one-stage d.c. amplifier. Its control grid receives the voltage from the radioactive collector via a voltage divider. One pole of the recorder is connected to the plate of the tube and the other one to the battery. Some points ought to be mentioned which are not always taken into proper consideration. The input resistance of the instrument should be higher than the resistance between air and collector surface, so that the input resistance does not load the collector.

VOL. 2 1952

To prevent an overload of the tube the large voltage of the collector has to be reduced in such a way, that the control grid does not receive more than 1 v. This is done by means of a dropping resistor of $3 \times 10^{11}\,\Omega$. The grid leak resistor is constructed as a voltage divider which allows the sensitivity to be changed within a wide range. With this arrangement it is possible to cover practically all potential gradient variations under fine weather conditions as well as during periods of extremely high potential gradients (showers, hail storms, thunderstorms) (compare K. BURKHART [2]).

For getting the diurnal variations, hourly means of the BENNDORF registration are read and again reproduced on a far smaller time scale. Since the diurnal variations are of utmost importance it seems to be justified to construct the intrument in such a way that it records the diurnal variations automatically in a convenient manner. The short fluctuations are of no interest and can be filtered away by increasing the time constant; that is, by connecting a capacitance of low loss across the grid leak. A time constant of 1 hr can easily be achieved because of the large resistance of the grid leak. This procedure also allows the length of the paper transport to be reduced down to 24 cm per day or even less without losing essential details.

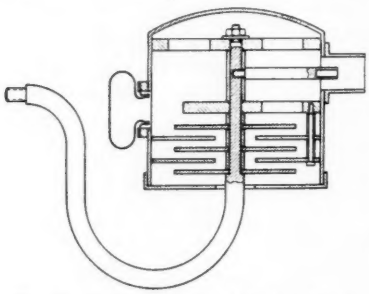


Fig. 3. Section of outside insulator.

To make readings linear over the whole scale it is essential to let the tube work on the straight part of the characteristic. To this end one can use mechanical or electrical suppression of the zero point of the recorder. Bridge or compensation circuits would give the same results but are avoided, since additional batteries and heavier battery drain would make the apparatus less convenient for use on expeditions.

The potentiometer P (Figure 2) permits the selection of the grid bias and, as a consequence of this, of the plate current, and allows the zero reading of potential gradient to be placed on any desired value of the recorder scale.

The most critical point of the arrangement is the maintenance of the high insulation resistance which must be at least a hundred times larger than the grid leak and dropping resistor together. Low loss cable as is used for coaxial lines in high frequency technique can be employed to connect the collector to the input jack of the amplifier. The insulation of the cable is not influenced by humidity and it was found that a length of 60 ft did not cause any trouble. Thus the attention of the observer has only to be concentrated on the connections. In a room where the relative humidity can always be kept well below 100 % difficulties arise only seldom. Quite different are the conditions at the outside insulator that supports the collector, a section of which is shown in Figure 3. It is constructed on the lines of an air insulated condenser. Its upper room is dried by means of calcium chloride which is kept in the glass vessel to the left of the insulator housing. The outside air has to pass the long way between the condenser plates before it is able to touch the insulating parts. The calcium chloride keeps the relative humidity below 100 % and has only seldom to be replaced, even during periods of wet weather. On mountain

peaks the observer has only to watch that snow and hoar frost crystals do not short-circuit the collector. At sea level spider's webs and small insects often spoil the insulation. A special protective kind of glue smeared on the insulator guards against insects. The spider's webs have to be removed in the evening before they become wet and lose their insulation.

III. Registration of Air-Earth Current

The difficulties in recording air-earth current do not arise only because of its small quantity (about 3×10^{-12} A/m^2 at sea level and during fine weather). Every receiver, usually a metal plate or net, exposed to it, not only receives the air earth-current but also the influence charges caused by the fluctuations of the potential gradient which give rise to currents 100 times larger than the air earth-current itself. On the direct current caused by the air-earth current is superposed an a.c. of variable frequency. The a.c. can be filtered away by means of a filter of convenient time constant, say 1 hr, which is also desirable from another point of view outlined above. It has been mentioned before that the range of the recorder is 0·1 mA. A plate current change of 0·05 mA (that is half the recorder range) of the electrometer tube can be caused by changing the grid bias by about 1 v, *i.e.* the grid leak resistor has to be so large that a current of 3×10^{-12} A causes a voltage drop of 1 v. It is possible to read currents down to 3×10^{-13} A from the registrations.

VOL. 2 1952

By means of a switch the grid leak resistance can be changed so as to keep the voltage variations of the control grid within 1 v, and to adapt the sensitivity of the amplifier to any occurring air-earth current. With regard to insulation and placing the electrical zero reading, the principles outlined in the previous chapter are valid.

The plate or net that picks up the air-earth current should always be at ground potential when placed level with the earth's surface [3]. The voltage drop across the grid leak resistor causes a potential difference between net and ground that varies with the air-earth current between 0 and 1 v. One could try to lift the receiving plate in such a way that its average potential is equal to that of the surrounding air. Under normal conditions the potential gradient is about 1 v/cm, *i.e.*, the receiving plane has to be lifted 0·5 cm. Such an adjustment is difficult, and besides experiments show that small differences in height and potential do not influence the results appreciably. The situation is quite different when the air-earth current becomes so large (*e.g.* during showers) that the receiving plate acquires a potential difference of 100 v or more above ground. This occurs when the observer does not switch down the sensitivity in time. If this happens the receiver catches more radioactive particles than usual (receiving plate on ground potential). When the air-earth current returns to normal values the radioactive material on the net causes a higher conductivity and in consequence of this a higher air-earth—current than the normal one for a period of about 2 hr; *i.e.*, until the radioactive material has disappeared. This fact has to be taken into consideration when taking readings after periods of bad weather during which the sensitivity has not been switched down.

IV. Registrations on the Nebelhorn, on the Zugspitze and at Buchau

A first test of the instrument was made in September 1949 on the Nebelhorn (7333ft). This test as well as that on the Zugspitze in March and April 1950 (9711 ft), was made to discover technical defects but the results yielded during this work are

already of some general interest and give some hints as to what we can expect from the planned mountain programme [4]. The apparatus was housed in a primitive canvas shelter on the western ridge close to the summit of the Nebelhorn. After ten days of heavy rain during which the amplifier failed to work properly due to poor insulation (the instrument was then in its first stage of development) a period of five days of Foehn followed. The air-earth current was found to be about ten times higher than at sea level. The potential gradient, which was extremely disturbed by the sharp ridge and therefore expected to be very high, amounted only to slightly below 100 v/m. The daily variations in air-earth current were steady throughout these five days and resemble very much the world time variations of potential gradient found at sea (Figure 4) except for the maximum at 4 hr GMT which is, rather surprisingly, missing in the registrations of potential gradient. The similarity to the world time variations indicates the absence of serious convection. Moreover, at greater heights one has to expect the stationary state to be reached in a shorter period of time than at sea level, due to the higher conductivity. Although these facts are quite striking it is not advisable to derive further conclusions from this small number of observations.

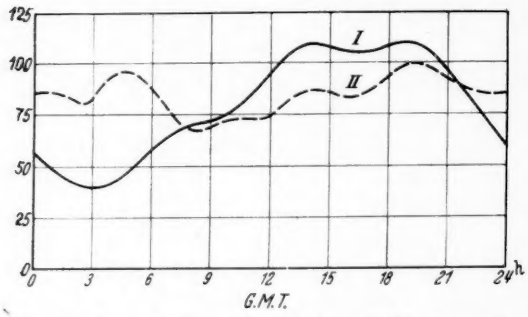


Fig. 4. Curve I: Diurnal variation of the air-earth current on the Nebelhorn. Curve II: Diurnal variation of the thunderstorm activity (F. J. W. WIPPLE and F. J. SCRASE).

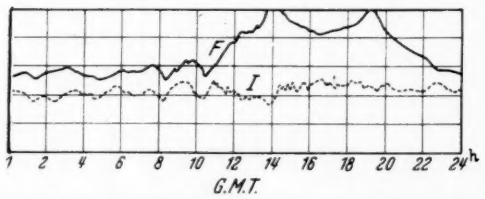


Fig. 5. Potential gradient (F) and air-earth current (I) of a typical day on the Zugspitze.

The registrations on the Zugspitze were started in the beginning of March 1950 in a period of high atmospheric pressure. P. LAUTNER [5] had previously carried out observations of potential gradient and air-earth current for three years on this mountain. The results differ considerably from those obtained on the Nebelhorn since no Foehn occurred during the observations. Figure 5 shows the variations of a typical day. The air-earth current does not have any apparent diurnal variation. The potential gradient has two maxima in the afternoon, namely one at 14 hr and the other one at 19 hr GMT. This result is well in accordance with the results yielded by LAUTNER on the Zugspitze and by other observers on high stations. The single period which is found on undisturbed days is most probably caused by sun radiation, which gives rise to convection at noon; that is, it lowers the conductivity and therefore lets the potential gradient rise. The air-earth current is not seriously affected by convection and therefore promises to throw more light on any world time variation than any other element. During the night and the morning hours potential gradient and air-earth current run fairly parallel. This fact demonstrates the absence of convection. For statistical investigations the material yielded on the Zugspitze is too poor. The potential

gradient variations observed at Buchau during the same time differ very much from those found on the Zugspitze due to local influences. In air-earth current the differences are smaller. As soon as more material is available statistical investigations will be started which should reveal important facts on the mechanism of the atmospheric electric field.

The registrations on the Zugspitze as well as those at Buchau seem to demonstrate that during undisturbed days air-earth current and potential gradient run parallel, especially during the night and the early hours of the morning. The parallelism is better on the Zugspitze than at Buchau and is upset as soon as convection starts. Since the material on which this conclusions is based is still poor it should be considered with caution; but it seems to show that the method proposed by H. ISRAËL (*loc.cit.* p. 32), *i.e.* to study convection by means of measuring atmospheric electric elements, promises some success.

Thanks are due to Dr. HANS ISRAËL for much helpful discussion and encouragement.

REFERENCES

[1] ISRAËL, H.; Die Tagesvariationen des elektrischen Widerstandes der Atmosphäre. Ann. Géophys. 1949 **5** 196–210. ISRAËL, H.; Probleme und Aufgaben der Luftelektrizität. Ber. dtsch. Wetterd. d. US-Zone 1950 (No. 12) 75–77. [2] BURKHART, K.; Luftelektrische Feldmessungen mit Elektrometerröhren. Z. Met. 1947 **1** 212–213. [3] WILSON, C. T. R.; On the measurement of the atmospheric electric potential gradient and the air-earth current. Proc. Roy. Soc. A 1908 **80** 537–547. [4] ISRAËL, H. see Nr. 1. [5] LAUTNER, P.; Deutschcs Meteorologisches Jahrbuch. Bayern, 1926–28, 1929–30.

VOL. 2 1952

Journal of Atmospheric and Terrestrial Physics, 1951, Vol. 2, pp. 38 to 50. Pergamon Press Ltd., London

Comparaison des résultats de mesures de l'absorption ionosphérique effectuées par deux stations européennes

K. RAWER

Service de Prévision Ionosphérique Militaire (France)

(Received 20 May 1951)

ABSTRACT

1. A general discussion of the usual method of absorption measurements reveals some difficulties of physical interpretation.

2. With the absorption measurements of Slough the following three results have been obtained: The seasonal variation follows a law with $\cos \chi^{\frac{3}{2}}$, but only in summer (Fig. 2). The influence of the solar cycle is important (Fig. 3). A considerable fluctuation exists from one day to another.

3. Such daily values measured at Slough and Freiburg (distance 400 km) have been compared (Fig. 4 and 5). They mostly show a positive correlation, but the correlation coefficient is only of the order of 0·4. Some of the difference seems to be brought about by local phenomena, but this conclusion is not yet certain, because the method of measurement is somewhat different.

A new measure of correlation, which is rapidly calculated, is proposed in the appendix.

1. LA MÉTHODE DE MESURE

Utilisée pour la mesure de la hauteur de réflexion, la méthode des impulsions a également été employée presque dès l'origine pour la mesure de l'absorption ionosphérique. Néanmoins le nombre des stations qui effectuent régulièrement ces mesures est ancore très réduit.

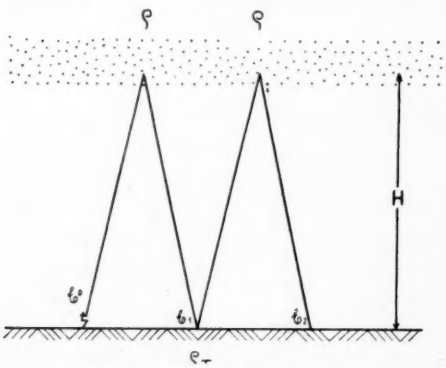


Fig. 1. Mesure du coéfficient apparent de réflexion.
Measurement of the apparent reflection coefficient.

La comparaison d'échos a été proposée par RATCLIFFE [1]. On peut deduire le coéfficient de réflexion de l'ionosphère du rapport des amplitudes de deux échos nés l'un d'une réflexion simple, l'autre d'une réflexion double sur l'ionosphère. Plus exactement, on obtient le produit de ce coefficient ϱ par le coefficient de réflexion du sol ϱ_T [2]. Une inconnue du problème, à savoir la puissance rayonnée par l'émetteur en direction normale, est ainsi éliminée. (Elle sera caractérisée ci-dessous par le champ E_0 à un km au dessus de l'antenne d'émission.) La hauteur de réflexion H ne s'introduit pas dans le calcul. Ceci est rapidement vérifié à l'aide des formules suivantes (Fig. 1).

$$E_1 = \frac{1}{2H}\,\varrho\,E_0; \qquad (1a) \qquad E_2 = \frac{1}{4H}\,\varrho^2\varrho_T E_0 \qquad (1b)$$

d'où

$$E_2 : E_1 = \frac{\varrho\,\varrho_T}{2}; \qquad (2a) \qquad E_0 = H\,\varrho_T E_1^2/E_2. \qquad (2b)$$

Ce calcul très simple est valable lorsque la terre et l'ionosphère sont considérées comme deux miroirs plans et parallèles. Il n'y a aucun intéret pratique à introduire ici (en incidence normale) les formules plus compliquées qui résultent de l'hypothèse d'une ionosphère et d'une terre sphériques et qui ne prennent toute leur importance qu'en incidence oblique [3].

En introduisant les logarithmes du champ, nous pouvons définir un décrement d'absorption δ_a qui s'écrit en décibels:

$$\delta_a = -20 \log \varrho = \langle E_1 \rangle - \langle E_2 \rangle - 6\,[db] + 20 \log \varrho_T . \tag{3}$$

(Les crochets $\langle\ \rangle$ indiquent que les champs en question sont mesurés en décibels.)

En général on admet que ϱ_T, le coefficient de réflexion du sol, est très voisin de l'unité; en conséquence on adoptera pour valeur de ϱ_T l'unité. Nous pouvons estimer sa valeur, par examen de réflexions multiples (triple, quadruple etc.) qu'on peut observer pendant la unit, où l'absorption ionosphérique est généralement très faible. Dans les stations sur mer, on peut observer les échos successifs en provenance de la couche F_2 jusqu'au dixième au moins. La perte due à une seule réflexion sur mer ne peut donc pas dépasser 1 décibel. Pour une station terrestre la perte est certainement plus grande; nous estimons 1 à 2 décibels pour Fribourg, or l'on n'observe 5 ou 6 échos au moins pendant la nuit (puissance de créte 1 à 2 kw). Notre station est bien située à ce point de vue, il existe certainement d'autres stations où l'on doit admettre une perte plus grande. Il nous semble que l'implantation d'une station ionosphérique doive donc se faire sur un terrain plan et de bonne conductibilité, (humide) dont le diamètre soit de 10 km au moins (ainsi plusieures des zones annulaires de FRESNEL se trouvent dans la région de bonne réflexion). Etant donné que les mesures nous donnent toujours le produit $\varrho\,\varrho_T$ il est à craindre que l'absorption soit surestimée, si la perte introduite par le sol est négligée.

VOL. 2 1952

Si les amplitudes des échos ont pu être observées correctement le décrement δ_a est obtenu par la formule (3). Mais dans les mesures de midi, surtout sur les fréquences rélativement basses, le deuxième écho est souvent trop absorbé, on ne peut alors donner qu'une limite supérieure pour son amplitude. Dans ces cas on déduit:

$$\delta_a = -20 \times \log \varrho = \langle E_0 \rangle - \langle E_1 \rangle - 20\,.\log\,(2H/[\mathrm{km}]) \tag{4}$$

la hauteur apparente de réflexion H doit donc être connue; elle peut facilement être mesurée.

En réalité nous n'appliquons la formule (4) que pour des échos en provenance de la couche E ou E_s, dont la hauteur est presque constante. Introduisant la valeur de 110 km nous avons avec une précision qui semble suffisante:

$$\delta_a = \langle E_0 \rangle - \langle E_1 \rangle - 46{\cdot}8\,[db] \tag{4a}$$

E_0, caractérisant le rayonnement de l'émetteur en direction normale, peut être déduit à l'aide de (2b) des observations antérieures ayant donné deux échos; (à la rigueur ont fait des observations pendant les heures de faible altitude du soleil). Naturellement il reste à vérifier que la puissance rayonnée n'a pas changé avec le temps. En réalité cette restriction n'est pas trop dangereuse (si la puissance changeait dans le rapport de 1 à 2 la valeur de E_0 ne varierait que de 3 décibels).

L'influence de la non consideration de ϱ_T existe aussi dans ce procédé malgré le fait que ϱ_T n'apparaisse pas dans la formule (4). C'est par E_0 qu'il s'introduit. Des mesures $E'_0 = H \times E_1^2/E_2^2$ est directement déduit et nous avons $E_0 = \varrho_T \times E'_0$, donc suivant (4)

$$\delta_a = -20 \times \log \varrho = \langle E'_0 \rangle - \langle E_1 \rangle - 20 \times \log\left(\frac{2H}{[\mathrm{km}]}\right) + 20 \log \varrho_T . \tag{4'}$$

En présence d'une perte d'énergie due au sol, les décrements d'absorption contiennent toujours cette influence indépendamment du procédé d'évaluation qui est appliqué. Il serait plus correct

d'y appliquer une correction. Celle-ci est plus importante pour les fréquences élevées où déja l'exactitude rélative des mesures est moins bonne.

La difficulté la plus importante rencontrée dans l'exécution des mesures est due aux fadings, que nous avons négligés dans nos formules. Certainement il serait imprudent de se contenter de la lecture d'une seule amplitude. Les variations rapides du champ dues aux variations de la polarisation sont éliminées si au lieu d'utiliser une seule lecture instantanée, on en recherche la médiane pendant au moins 20 sec. Mais il y a en général aussi des variations lentes de l'ordre de plusieures minutes. Parfois même l'ordre habituel des amplitudes est inversé, le deuxième écho est alors plus fort que le premier. Naturellement nos formules ne peuvent pas être utilisées dans un tel cas (si non, on obtiendrait une valeur négative du décrement d'absorption). Ces variations lentes ne peuvent guère s'expliquer par des phénomènes d'interférence. Vraisemblablement il s'agit d'ondulations dans les surfaces d'égale ionisation [4]. Dans ces cas, même si l'on constate que le premier écho est d'amplitude supérieure a celle du second, on peut en attribuer la cause non seulement à une perte par absorption mais aussi à un effect de défocalisation. C'est pourquoi la comparaison de la médiane des amplitudes obtenue au cours d'une minute ne suffit pas encore.

Afin que de telles ereurs soient évitées dans toute la mesure du possible, une mesure d'absorption devrait s'étendre sur au moins 15 min et les observations devraient toujours être effectuées sur plusieures fréquences [5]. Pour une fréquence fixe (c'est à dire dans un certain niveau de l'ionosphère) un effet de focalisation ou de défocalisation peut se maintenir plus longtemps; mais il est peu probable qu'il se maintienne dans toute la gamme des fréquences.

Pour en tenir compte nous avons adopté à Fribourg le procédé suivant*. L'émetteur fonctionne en général avec variation continue de la fréquence; il peut au moyen d'un simple interrupteur être très facilement manipulé en fontionnement discontinu [6]. Dans ces cas l'émission a lieu sur 8 fréquences fixes (1·4, 2·0, 2·8, 4·0, 5·6, 8·0, 11·2, 16·0 MHz) dont quatre sont en général notées (2·0 . . . 5·6 MHz). L'émission de chaque fréquence dure 75 sec. Après le premier balayage des fréquences qui dure 10 min, l'observation sur ces 4 fréquences est renouvelée de sorte que la durée totale des observations est de 20 min. Chacune des 4 fréquences est donc observée deux fois dans un intervalle de 10 min. Pendant chaque période d'observation (60 sec) l'amplitude des échos visibles sur l'écran d'un tube à rayon cathodique est notée 7 fois, les lectures ayant lieu à des intervalles de 10 sec. Leur médiane et celle de la répétition suivante (10 min après) sont moyennées. Ainsi obtient on les valeurs de l'amplitude qui sont utilisées pour le calcul de δ_a.

Observer pendant 20 min nous semble suffisant. Nous devons admettre qu'il existe encore au cours des 20 min d'observation des cas de focalisation. D'autre part, l'absorption elle même subit probablement de rapides variations. Plus l'on observe longtemps, plus l'on risque de voir s'estomper l'effet de ces variations, ce qui serait facheux. La durée de 20 min nous semble donc être un compromis tolérable.

Les mesures sont faites sur plusieures fréquences. Les résultats sont présentés en général par la formule:

$$\delta_a = A/(f + f_L)^2 \tag{5}$$

valable seulement pour une absorption non sélective comme celle de la couche D. (Approximation quasilongitudinale, rayon ordinnaire [7]; f_L est la composante longitudinale de la gyrofréquence, en Europe $f_L \cong 1{\cdot}2$ MHz.) La valeur A obtenue

* Avant 1950 un procédé sémblable mais dont l'horaire était moins précis a été suivi.

ainsi est supposée caractériser l'absorption de la couche plus basse de l'ionosphère, à savoir la couche D; son unité est [décibels MHz^2]. Il y a encore une certaine influence de l'absorption sélective due è la couche E; cette question sera traitée dans une autre publication [8].

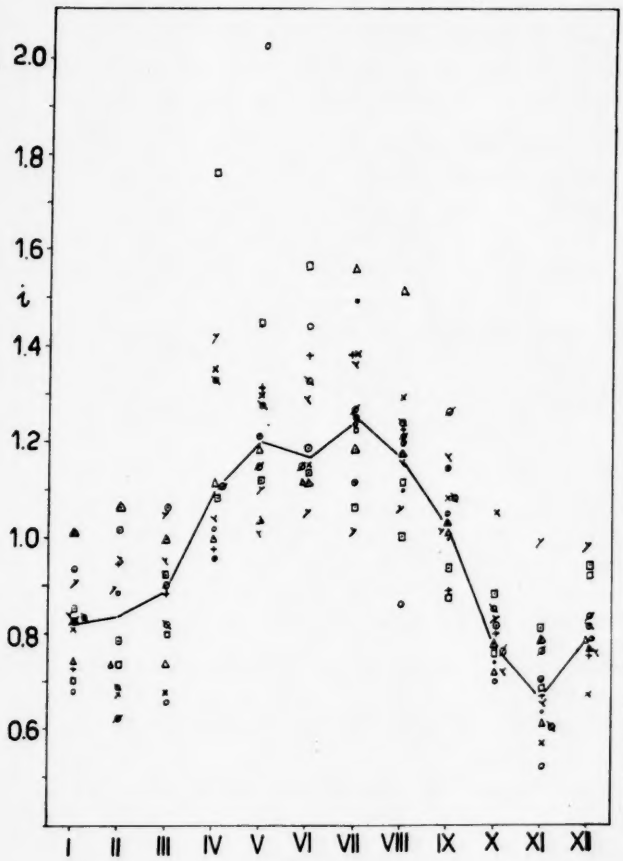


Fig. 2a.

2. Analyse Statistique des Mesures Anglaises

Sur l'initiative de Sir Edward Appleton [9] la station anglaise de Slough a effectué régulièrement des mesures d'absorption à partir de 1934. Nous avons entrepris avec son accord

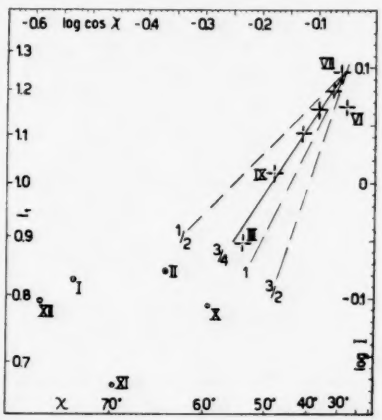


Fig. 2b.

Fig. 2a. Variation saisonnière des valeurs de l'absorption mesurées à midi (Slough 1934–1947). $i = \bar{A}/\bar{A}$; I = méd (i). Seasonal influence on absorption at noon (Slough 1934–1947).

Fig. 2b. Variation saisonnière comparé à l'angle solaire zénithal χ. (Echelles logarithmiques, les droites correspondant à différentes valeurs de l'exposant de cos χ; les croix indiquent les mesures d'été, les cercles celles d'hiver.)

Seasonal variation versus the zenith-angle of the sun, χ. [Logarithmic scales; the lines indicate different exponents of cos χ; measurements are represented by crosses (summer) and by circles (winter).]

une analyse statistique de cette série très précieuse dont les résultats * [10] peuvent être résumés comme suit**.

1. Les médianes mensuelles des valeurs de l'absorption mesurées à midi, $\bar{A}$, mettent nettement en évidence l'influence saisonnière (Fig. 2a). En été la

* Les résultats de ces calculs confirment les conclusions de Sir Edward Appleton et W. R. Piggot, URSI 1948 **7** p. 320 No. 171.

** Published by permission of the Director of Radio Research, Department of Scientific and Industrial Research, London. Crown Copyright reserved.

variation est conforme à un loi de $\cos^{3\frac{1}{2}}\chi$ (χ la distance zénithale du soleil), tandis que l'influence de la hauteur du soleil semble disparaitre en hiver (Fig. 2b). Appleton a été le premier à remarquer que la rapport de l'absorption en été et en hiver correspond à une autre loi que la variation journalière [9].

2. Il existe une influence remarquable du cycle undécennal du soleil; elle peut être démontré, à l'aide des moyennes glissantes, $\overline{\overline{A}}$, calculées sur 12 mois (Fig. 3). La relation avec la moyenne glissante du nombre relatif des tâches solaires $\overline{\overline{R}}$ est

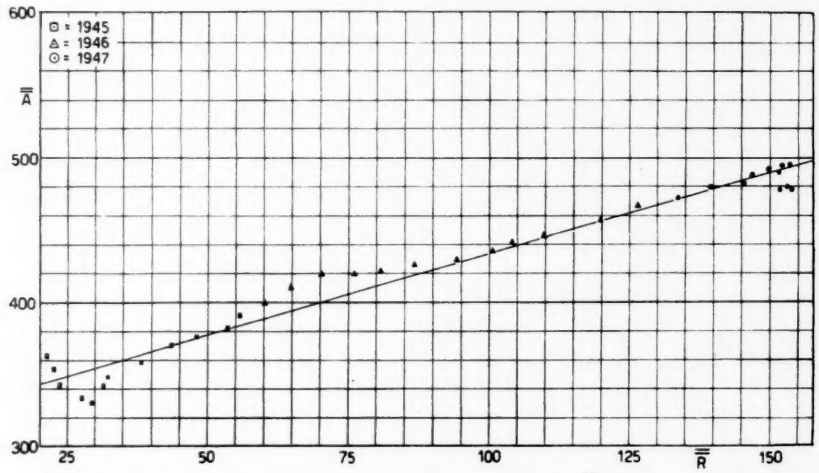


Fig. 3. Correlogramme entre la moyenne glissante de l'absorption $\overline{\overline{A}}$ et celle du nombre des taches solaires $\overline{\overline{R}}$.

Correlogram of sliding means of the absorption $\overline{\overline{A}}$ with those of the Wolf-number $\overline{\overline{R}}$.

presque linéaire, le coefficient de corrélation est très élevé à savoir 0·98. Nous obtenons

$$\overline{\overline{A}} = 320\,[db\mathrm{MHz}^2] \times (1 + 0{\cdot}0035\,\overline{\overline{R}}) \qquad (2)$$

Cette relation a été déduite des mesures de la période de croissance du cycle undécennal de 1945 à 1948. On remarque sur la Fig. 3 que l'echelonnement des premiers points indiquait une diminution. Ceci semble du au fait que le minimum de l'absorption ne coincide pas exactement avec celui des taches solaires.

3. Les valeurs journalières de l'absorption ont une dispersion considérable. Les caractéristiques de la couche D sont beaucoup plus variables que celles de l'ionisation de la couche E normale.

3. Comparaison des Mesures de Slough et de Fribourg

La station du SPIM à Fribourg (Allemagne) a commencé des mesures comparables en 1948 [11]. Il est fort probable que les influences de la saison et du cycle solaire sont assez voisins de ce qui a été observé à Slough, mais il nous a semblé intéressant de comparer les mesures journalières de ces deux stations distantes de 400 km. On pourrait même s'attendre à en tirer une conclusion sur la cause des fluctuations

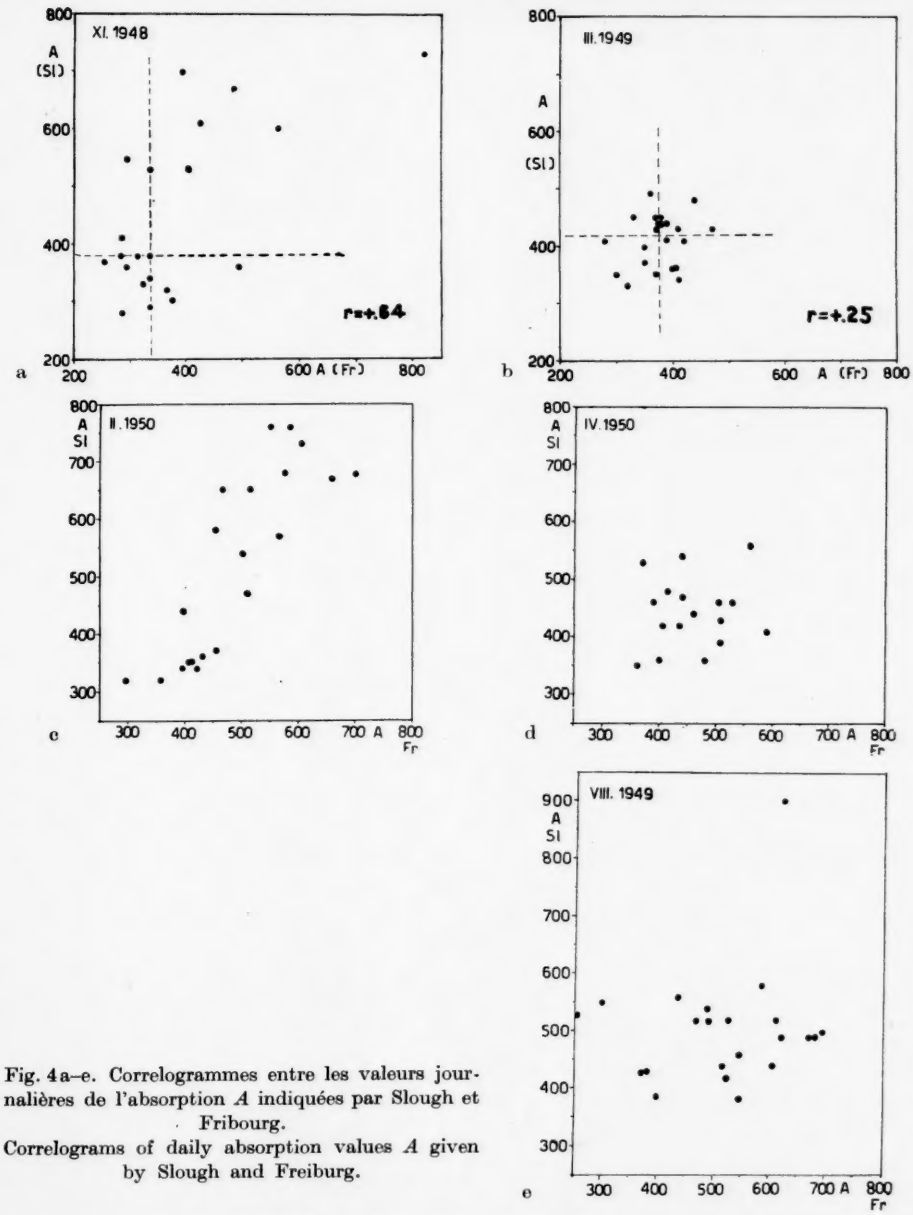


Fig. 4a–e. Correlogrammes entre les valeurs journalières de l'absorption A indiquées par Slough et Fribourg.
Correlograms of daily absorption values A given by Slough and Freiburg.

de l'absorption. S'il s'agissait d'une influence purement solaire les variations devrait être synchrones.

Nous avons établi les corrélogrammes mensuels de la valeur A pour 26 mois de 1948 à 1950 (chaque point correspond à une mesure). Les Fig. 4 en montrent quelques exemples. Il nous a paru intéressant d'estimer la corrélation d'après une méthode moins laborieuse que l'est le calcul du "coefficient de corrélation." Nous

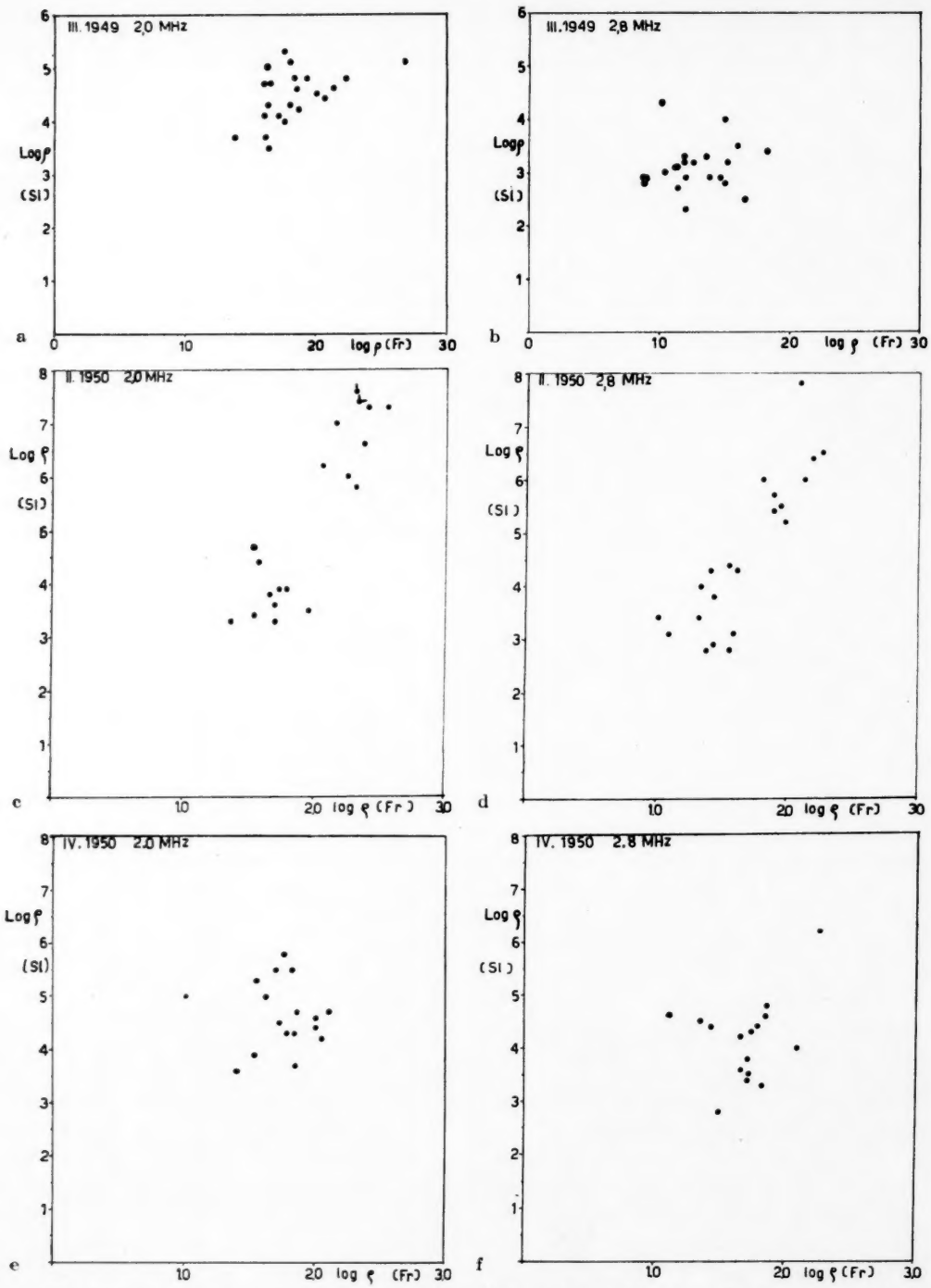


Fig. 5a–f. Correlogrammes entre les coefficients d'absorption ϱ obtenus sur les fréquences 2 et 2·8 MHz d'après les mesures journalières de Slough et de Fribourg.
Correlograms of the reflection coefficient ϱ obtained on the frequencies 2 resp. 2·8 MHz at Slough and Freiburg.

proposons, en annexe, un autre terme nommé "nombre de corrélation" dont la valeur est très rapidement obtenue à l'aide du corrélogramme. Pour une partie de la periode considérée, les deux mesures ont éte évaluées à titre de comparaison. Les résultats sont donnés dans le Tableau 1.

Pour la comparaison de 26 mois nous obtenons 8 fois un nombre de corrélation négatif; les valeurs obtenus sont comprises entre $-0{\cdot}80$ et $+0{\cdot}97$. La médiane des

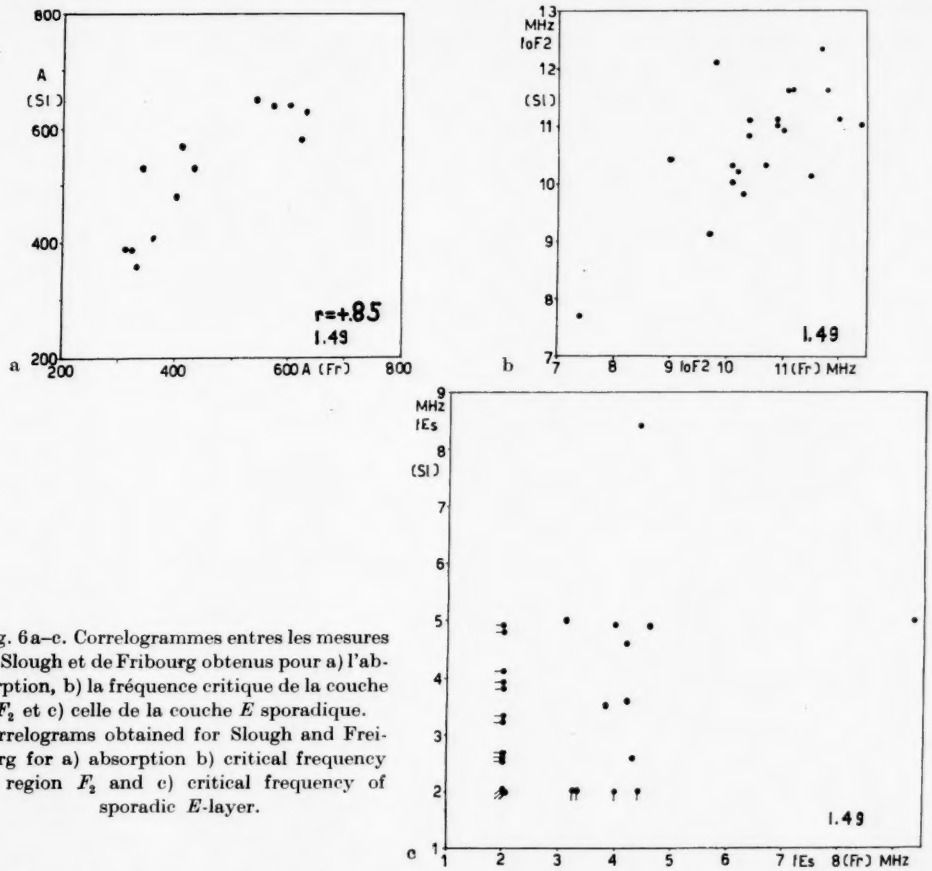


Fig. 6a–c. Correlogrammes entres les mesures de Slough et de Fribourg obtenus pour a) l'absorption, b) la fréquence critique de la couche F_2 et c) celle de la couche E sporadique. Correlograms obtained for Slough and Freiburg for a) absorption b) critical frequency of region F_2 and c) critical frequency of sporadic E-layer.

26 nombres R est seulement $+0{\cdot}33$. Dans ces circonstances l'existence d'une corrélation physique entre les mesures journalières des deux stations semble assez douteuse. Notons encore que les valeur négatives de R n'ont été obtenues qu'en été. Ceci est peut-être à attribuer à l'influence de E sporadique (?)

D'autre part les Fig. 5 et le Tableau 2 sont à base des valeurs du décrement obtenu directement sur certaines fréquences observées par les deux stations. Nous obtenons ainsi un résultat plus encourageant; la médiane des 26 mois devient 0·5 au lieu de 0·33. Néanmoins il nous reste encore plusieurs cas d'une corrélation mauvaise et même négative.

Il est intéressant de comparer ce résultat avec les mesures des autres couches ionosphériques (Fig. 6 en donne un exemple pour un mois où les valeurs d'absorption

VOL. 2 1952

Tableau 1. Corrélation des valeurs A
(r = coéfficient de corrélation; R = nombre de corrélation)

	I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*	*IX*	*X*	*XI*	*XII*
1948												
r	—	—	—	—	—	—	—	—	—	—	—	—
R	—	—	—	—	—	—	—	—	—	—	0·47	0·63
1949												
r	0·85	0·60	0·25	0·26	0·90	−0·30	0·39	+0·27	−0·19	0·17	0·68	0·70
R	0·97	0·74	0·18	0·45	0·60	−0·67	0·28	−0·35	−0·25	−0·80	0·73	0·60
1950												
r	0·64	0·82	0·36	+0·22	0·14	−0·15	(−0·08)	−0·50	—	—	—	—
R	0·22	0·97	0 09	−0·11	0·38	−0·60	(+0·09)	−0·44	(−0 06)	0·69	0·82	0·72

médiane des valeurs R: + 0·33

Tableau 2. "Nombre de corrélation" des décrements δ_a mesurés sur deux fréquences

	I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*	*IX*	*X*	*XI*	*XII*
1948												
2·0 MHz . .	—	—	—	—	—	—	—	—	—	—	0·53	0·85
2·8 MHz . .	—	—	—	—	—	—	—	—	—	—	0·77	0·60
1949												
2·0 MHz . .	0·77	0·77	0·51	−0·10	0·22	0·34	0·18	0·22	−0·17	0·47	0·82	0·97
2·8 MHz . .	0·96	0·75	0·35	0·26	0·38	0·19	0·43	0·35	−0·18	0	0·87	0·50
1950												
2·0 MHz . .	−0·22	0·99	0·68	−0·34	0·55	−0·22	−0·18	0·06	(0·32)	0·65	0·85	0·81
2·8 MHz . .	0·77	0·96	0	0·47	−0·47	−0·18	0·35	0·44	(−0·60)	−0·22	0·13	0·00

médiane des valeurs R: *0·50* (2·0 MHz) *0·35* (2·8 MHz)

avaient une bonne corrélation). Tandis que l'ionisation des couches F_2 et E subit à peu près les mêmes variations pour deux stations distantes de 400 km [12], le coëfficient de corrélation est très faible pour la couche E_s [13]. Nous en tirons la conclusion, que la couche D est à peine plus homogène que la couche E_s.

Considérant les détails des mesures nous n'osons pas encore tirer rigoureusement cette conclusion. A fortiori nous ne pouvons pas établir définitivement notre opinion suivant laquelle les fluctuations de l'absorption ne sont pas dues à une influence purement solaire. Considérant la méthode de mesure des deux stations, nous trouvons quatre raisons en faveur d'un jugement prudent sur ces mesures.

(a) Les mesures ne sont pas exactement synchrones. A Fribourg elles sont commencées à peu près à 11^{40} TU; avant l'année 1950 la durée était à peu près 40 min maintenant elle est exactement 20 min (voir paragraphe 1). Slough indique "midi" comme heure des mesure (12^{00} TU), mais la période de mesure n'est pas forcément identique à la nôtre.

(b) Les méthodes d'observation des deux stations ne sont pas exactement identiques. Slough fait l'observation sur huit fréquences différentes, Fribourg en utilise effectivement quatre. D'après notre information les différents échos sont mesurés successivement à Slough, les lectures n'ont pas lieu au même instant. Un atténuateur dans un récepteur à amplification *linéaire* est employé pour les mesures. Par contre nous observons à Fribourg à l'aide d'un récepteur *logarithmique* et d'un oscilloscope gradué; en effet les lectures des différents échos peuvent ainsi être obtenues au même instant.

(c) La méthode de dépouillement n'est pas suffisamment précisée. Malgré la formule (1) qui est toujours utilisée, le choix d'une valeur définitive de A parmi les résultats des différentes fréquences suppose un certain jugement personnel; ceci est

Tableau 3. Médianes mensuelles de l'absorption $\bar{A}$*

	I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*	*IX*	*X*	*XI*	*XII*
1948												
Sl	—	—	—	—	—	—	—	—	—	—	380	450
Fr	—	—	—	—	—	—	—	—	—	—	330	360
1949												
Sl.	530	440	420	450	435	440	470	495	400	395	380	415
Fr	410	360	380	380	450	510	555	520	475	355	335	370
1950												
Sl.	355	540	380	440	420	420	(420)	460	(480)	375	400	470
Fr	325	465	335	440	480	430	485	570	425	275	370	350

médiane Slough: 430 médiane Fribourg: 395

* Nous avons tenu compte seulement des jours pour lesquelles les deux stations ont donné une valeur A.

VOL. 2 1952

important parce que la dispersion des valeurs est assez élevée. Nous voyons par comparaison des Tableaux 1 et 2 que la différence introduite par cette appréciation diminue considérablement la valeur de la corrélation. La standardisation d'une méthode de dépouillement serait très intéressante.

(d) Les caracteristiques de la couche D semblent subir de rapides variations, de sorte que les conditions ne sont souvent pas constantes au cours d'une mesure. D'autre part l'élimination des fadings à longue durée nécéssite une durée assez longue de celle-ci. Dans ces conditions, les valeurs mesurées ne représentent qu'une moyenne pour la durée de la mesure, mais elles ne peuvent certainement pas être regardées comme valables pour plusieures heures. Nous espérons d'ailleurs obtenir d'ici peu des résultats plus nets par l'emploi de deux sondeurs du même type et bien synchronisés, de sorte que les mêmes fréquences seront observées pendant les mêmes minutes.

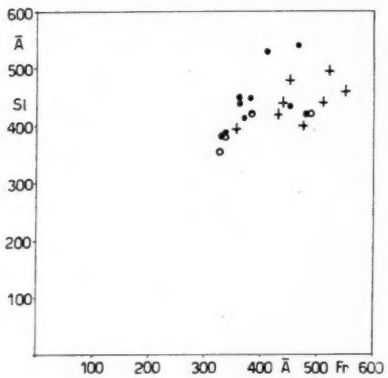


Fig. 7. Corrélogramme entre les médianes mensuelles $\bar{A}$ obtenues à Slough et à Fribourg. Les valeurs de septembre à decembre 1950 ont été ajoutées plus tard (Tableau 3); elle ne figurent pas ici. Correlogram of monthly median values $\bar{A}$ obtained at Slough and Freiburg.

Malgré ces difficultés la dispersion des 30 observations mensuelles présente un certain intérét et elle a certainement une signification physique. Nous trouvons qu'elle est à peu près la même à Slough et à Fribourg, soit $\pm 20\,\%$ pendant l'été et $\pm 30\,\%$ en hiver. Enfin la valeur absolue de la médiane mensuelle $\bar{A}$ est significative elle-aussi. Elle diffère un peu pour les deux stations (voir Tableau 3), et il semble que les valeurs de Fribourg soient en général un peu plus faibles. Mais la médiane des 26 mois d'observation, $\bar{A}$, est sensiblement la même. Nous avons établi le corrélogramme des $\bar{A}$ (Fig. 7) indiquant une bonne corrélation ($R = 0{\cdot}51$). Si nous éliminons les mois ayant donné un mauvais résultat dans le Tableau 1 (valeurs négatives de R, correspondant aux croix de la Fig. 7) nous obtenons même $R = 0{\cdot}72$. Nous avons l'impression qu'il soit plus difficile d'obtenir des résults comparables en été à cause de l'existence de la couche E_s.

Malgré toutes les difficultes que nous n'avons pas cachées il semble désirable que plus de stations ionosphériques entreprennent des observations régulières de l'absorption ionosphérique. C'est encore moyen le plus direct d'arriver à une étude suivie des propriétés de la basse ionosphère.

ANNEXE

Sur une méthode approchée permettant l'évaluation rapide de la corrélation existant entre deux grandeurs.

Pour de nombreuses applications statistiques, surtout en géophysique, on se sert maintenant plutôt de la valeur *médiane* que de la moyenne arithmétique.

Soient $x_1, x_2 \ldots x_n$ les mesures caractérisant un phénomène déterminé et telles que $x_1 < x_2 \ldots < x_n$. La médiane X des x_i est définie comme suit:

Pour n pair: $$X = \tfrac{1}{2}\,(x_{n/2} + x_{n/2+1})$$

Pour n impair: $X = x_{(n+1)/2}$.

Par définition nous avons les inégalités:

$x_i \leq X$ (pour la première moitié des valeurs x_i)

$x_i \geq X$ (pour l'autre moitié).

Dans les applications statistiques la médiane a un grand nombre d'avantages par rapport à la moyenne. Soit $f(x_i)$ une transformation monotone opérant sur les grandeurs x_i, la médiane subira la transformation* $f(X)$, ce qui n'est pas valable pour la moyenne. D'autre part elle n'est pratiquement pas influencée par la présence de valeurs aberrantes des x_i. Pour un nombre p de valeurs de x_i connues comme étant supérieures à une certaine limite imposée par l'expérience et tel que $p < \frac{1}{2}n$ la médiane peut encore être évaluée. Enfin la propriété de la médiane de correspondre à une probabilité de 50% est très intéressante pour les applications pratiques.

Soient maintenant x_i et y_i deux ensembles de n valeurs, dépendant du même paramètre t. On connait le calcul assez laborieux permettant de déduire le coefficient de corrélation. Au lieu de ce calcul appliquons un procédé graphique, basé sur le "corrélogramme" [champ des valeurs suivant les axes x et y, où chaque point permet de déterminer les valeurs de $x(t)$ et $y(t)$, correspondant à une même valeur t]. Tracons la médiane X des x_i et de même celle Y des y_i. Nous effectuons ainsi une division des points réprésentatifs en quatre parties A, B, C, D, telle que

$A)$ $x_i \leq X;\ y_i \geq Y.$ $\quad B)$ $x_i \geq X;\ y_i \geq Y.$

$C)$ $x_i \leq X;\ y_i \leq Y.$ $\quad D)$ $x_i \geq X;\ y_i \leq Y.$

(*A* titre d'exemple, cette division a été indiquée sur les Fig. 4a et b.) Si par hazard plusieurs x_i sont égals à X, on évaluera les fractions appartenant aux deux moitiés (c'est le cas de la Fig. 4a).

D'après la définition de la médiane:

$$A + C = B + D = \tfrac{1}{2}n; \quad A + B = C + D = \tfrac{1}{2}n$$

d'où

$$D = A; \qquad C = B = \tfrac{1}{2} \times n - A.$$

Considérons le rapport:

$$u = \frac{BC}{AD} = \left(\frac{B}{A}\right)^2 = \left(\frac{n}{2A} - 1\right)^2.$$

La valeur absolue de cette expression est grande, si la corrélation est bonne et si y_i est croissant en même temps que x_i; elle est petite dans le cas de décroissance. Elle est à peu près unité dans les cas où la corrélation est mauvaise.

Cette définition est "efficace" dans le sens de R. A. FISHER**. Introduisons maintenant par

$$R = \text{tgh} \log u = \frac{u-1}{u+1}$$

un "*nombre de corrélation*" qui peut varier entre -1 et $+1$. Notre nombre R se calcule beaucoup plus rapidement que le coefficient de corrélation. Il possède

* Cette expression est valable pour n impair et approchée pour n pair.

** FISHER, R. A.; Statistical Methods for Research Workers. Edinburgh.

VOL. 2 1952

deux des avantages de la médiane: 1. Il n'est pas influencé par quelques valeurs aberrantes. 2. Sa valeur absolue est invariante pour les transformations montones des x et y.

Nous avons comparé par de nombreux exemples la signification physique du "nombre de corrélation" R (d'après la définition donnée ci-dessus) et celle du "coefficient de corrélation" r (définition habituelle). Les deux nous conduisent à des résultats satisfaisants, si la reparatition des points du corrélogramme est normale; c'est à dire si les lignes d'égale probabilité sont des ellipses concentriques dont la valeur maximum se trouve au milieu. Dans ces cas l'ordre de grandeur du résultat est à peu près la même ($r \cong R$). Or si la repartition est irrégulière, surtout si le nombre des points est petit, il peut exister une différence appréciable entre R et r. Ce n'est pas forcément la mesure habituelle r qui s'adapte mieux à la réalité physique. [Le cas d'une fonction non-linéaire $y(x)$ conduisant à $r < 1$, malgré l'existence d'une corrélation physique exacte, est bien connu.] On peut trouver des exemples où soit r, soit R ne correspondent pas à la réalité physique. En général ce sont les points aberrants qui causent une différence entre r et R*. Leur influence sur r est grande tandis qu'ils n'ont aucune importance pour R par définition. Nous retrouvons ainsi la différence entre les caractéristiques de la moyenne et de la médiane.

Dans l'application physique on a parfois interêt de tenir compte des points aberrants; dans ces cas la moyenne et le coefficient de corrélation seront utiles. Mais parfois, surtout en géophysique, il est préférable que l'influence de ces points soit éliminée. Alors la médiane et une mesure similaire de la corrélation (p.e. le nombre R) correspondent mieux à la réalité. En conséquence notre procédé s'adapte bien à de nombreux problèmes géophysiques où la dispersion des points du corrélogramme est importante. Il a d'autre part l'avantage d'une évaluation rapide.

Dans le cas de nos exemples (Fig. 4a et b) nous obtenons ainsi:

$$n = 21;\ A = (3 + \tfrac{5}{6} + \tfrac{5}{16});\ u = (2{\cdot}67 - 1)^2 = 2{\cdot}79;\ R = 0{\cdot}47\ \ (r = 0{\cdot}64) \qquad (4a)$$

$$n = 22;\ A = 5;\ u = (\tfrac{8}{5})^2 = 1{\cdot}44;\ R = 0{\cdot}18\ \ (r = 0{\cdot}25). \qquad (4b)$$

BIBLIOGRAPHIE

[1] RATCLIFFE, J. A.; Proc. Roy. Soc. A 1935 **151** 370. [2] PIGGOT, W. R.; communication verbale. [3] RAWER, K.; Revue Scientifique 1948 **86** 481–485 et 585–600; Nature, Lond. 1950 **166** 316. [4] PIERCE, J. A. and MIMNO, H. R.; Phys. Rev. 1940 **57** 95–105. [5] COX, J. W.; The Measurement of Ionosperic Absorption, International Radio Propagation Conference, Washington (1944) (= IRPL-C18). [6] HARNISCHMACHER, E. et PHARABET, L.; Sondeur Ionosphérique KRDSE 22, rapport SPIM-C2. Service de Prévision Ionosphérique Militaire (1950). [7] BOOKER, H. G.; Proc. Roy. Soc. A 1935 **150** 267–286. [8] BIBL, K. and RAWER, K.; J. Atmosph. Terr. Phys. 1951 **2** 51. [9] APPLETON, E. V.; Proc. Roy. Soc. A 1937 **162** 451–479. [10] RAWER, K.; Les paramètres de l'absorption des ondes par la couche D, et leur prévision, rapport SPIM-R5, Service de Prévision Ionosphérique Marine (1949). [11] Bulletins SPIM-028-30 (mesures octobre–décembre 1948) et suite, Service de Prévision Ionosphérique Militaire. [12] BRAND, J. O.; Vergleich des Zustandes der F_2-Schicht an verschiedenen Orten, Schriften Dtsch. Akad. Luftfahrtforschung 1943 **7** 11. [13] HÖCHTL, F. et RAWER, K.; Ann. Géophys. 1949 **5** 150–156.

* Un seul point est décisif p.e. pour la valeur r en mai et en aout 1949 de notre Tableau 1 (voir Fig. 4e).

Journal of Atmospheric and Terrestrial Physics, 1951, Vol. 2, pp. 51 to 65. Pergamon Press Ltd., London

Les contributions des régions *D* et *E* dans les mesures de l'absorption ionosphérique

K. Bibl et K. Rawer

Service de Prévision Ionosphérique Militaire, France

(Received 1. July 1951)

ABSTRACT

Observations of ionospheric absorption made on different frequencies reveal an important contribution of selective absorption occurring in the *E*-layer. This influence can be calculated with the model of a parabolic *E*-layer and an exponential variation of the collision number. With the results of this calculation a new evaluation of the measurements of Slough and Freiburg has been made for 24 months. A graphical method is proposed, which gives separately the contributions of *E*- and *D*-layer. These values for *D*-absorption are, of course, not as high as can be obtained by the usual method, where the influence of the *E*-layer is neglected.

In spite of a considerable dispersion this evaluation leads to the adoption of a rather constant value of the influence of the *E*-layer. It is proposed to use this value for future reduction of absorption measurements. Finally this constant leads to a new determination of the number of collisions in the centre of the *E*-layer, which is certainly below 10^4 [s^{-1}] ($6{,}5\times10^3$ [s^{-1}]). This value, lower than those of other observers, is in good accordance with recent temperature and pressure data at 125 km.

VOL. 2 1952

1. Résultats des Mesures de l'Absorption

Dans un article précédent [1] un des auteurs a discuté la méthode des mesures de l'absorption ionosphérique et son interprétation habituelle. Dans celle-ci l'absorption est attribuée à la couche *D*, cause d'une absorption non selective, qui diminue rapidement avec la fréquence suivant la formule:

$$\delta_a = A/(f+f_L)^2. \tag{1}$$

Cette formule est valable pour le rayon ordinaire en approximation quasilongitudinale pourvu que la fréquence f ne soit pas trop basse (δ_a est le décrément d'absorption observé sur la fréquence f; f_L la composante longitudinale de la gyrofrequence, $f_L \approx 1{,}2$ MHz). La grandeur A, caractérisant l'absorption, devrait être indépendante de la fréquence; en général A est donné en [decibels MHz^2].

Cette formule peut être vérifiée si l'on dispose d'observations sur plusieures fréquences. La graphique des δ_a en fonction de $(f+f_L)^{-2}$ devrait être une ligne droite, dont la pente serait égale à A.

Or nous avons appliqué ce procédé aux mesures de Slough (Angleterre)* et à celles de la station du SPIM à Fribourg (Allemagne). Les mesures ont été effectuées chaque midi. Nous présentons les resultats des premiers six mois de l'année 1950 dans les Fig. 1 et 2. On y voit que les décrements ont tendance à diminuer avec la fréquence, mais on constate de même que le lieu des points n'est pas une ligne droite. Aussi nous avons une bosse entre 3 et 4 MHz qui semble avoir son maximum à la fréquence critique (f_0E), de la couche *E*. C'est pourquoi nous sommes amenés

* Published by permission of the Director of Radio Research, Department of Scientific and Industrial Research, London. Crown Copyright reserved.

à attribuer l'irrégularité mentionnée à l'influence d'une *absorption sélective*, ayant lieu près du maximum d'ionisation de la couche *E*.

D'ailleurs les courbes journalières sont beaucoup moins régulières, que les courbes médianes mensuelles. (Les médianes de dix jours montrent encore une dispersion considérable.) Les raisons de ce manque d'homogénite ont été discutées dans l'article précédent. Nous sommes arrivés enfin a la conclusion que la moyenne statistique d'un mois a certainement une signification physique. Les médianes

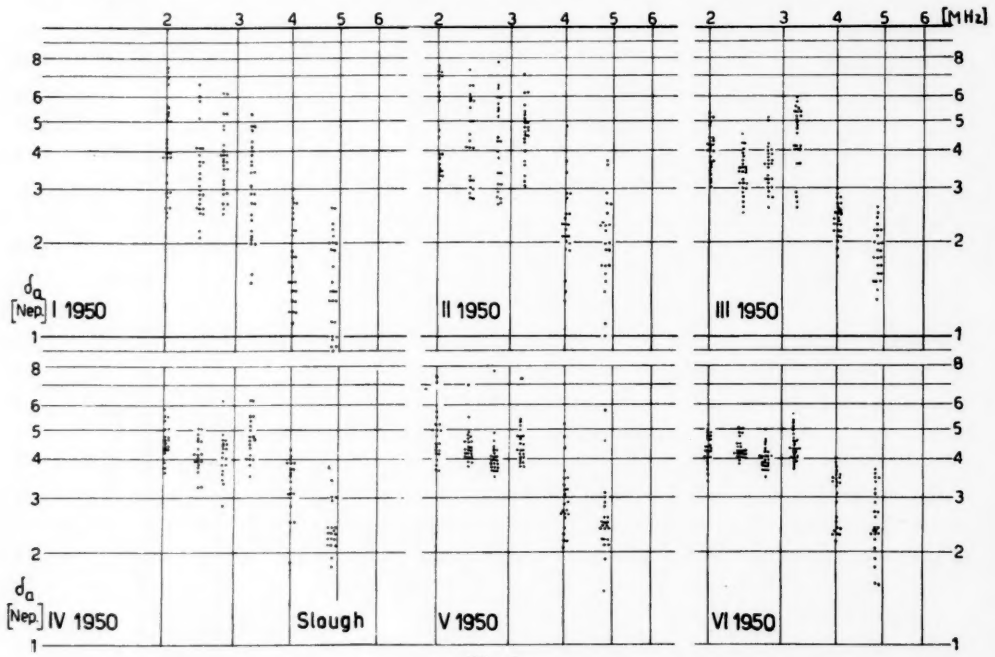


Fig. 1.

Fig. 1 et 2. Décrement de l'absorption en
Observed decrements of absorption

mensuelles de la valeur du décrement sont tabulées pour les stations de Slough et Fribourg dans les tables 1 et 2. Ces chiffres sont la base de l'analyse suivante. Les méthodes selon lesquelles le décrement δ_a a été obtenu figurent en détail dans l'article précédent.

2. Calcul de l'Absorption Sélective

Ce type d'absorption est seulement important dans une certaine gamme de fréquences entourant la fréquence critique de la couche *E*. L'absorption est proportionelle à l'épaisseur de la couche et au nombre de chocs (au niveau du maximum d'ionisation). Etant donné que l'absorption sélective a lieu à proximité du point de réflexion l'approximation quasitransversale est ici valable [2]; l'expression du coéfficient d'absorption k (rayon ordinaire) est donc:

$$k = \frac{\nu}{2c} \times \frac{1}{\mu} \times f_0^2/f^2. \qquad (2)$$

(ν nombre de chocs, μ indice de refraction, f_0^2 est proportionnel à la densité électronique N_e). Le décrement d'absorption sélective s'obtient par intégration lelong du

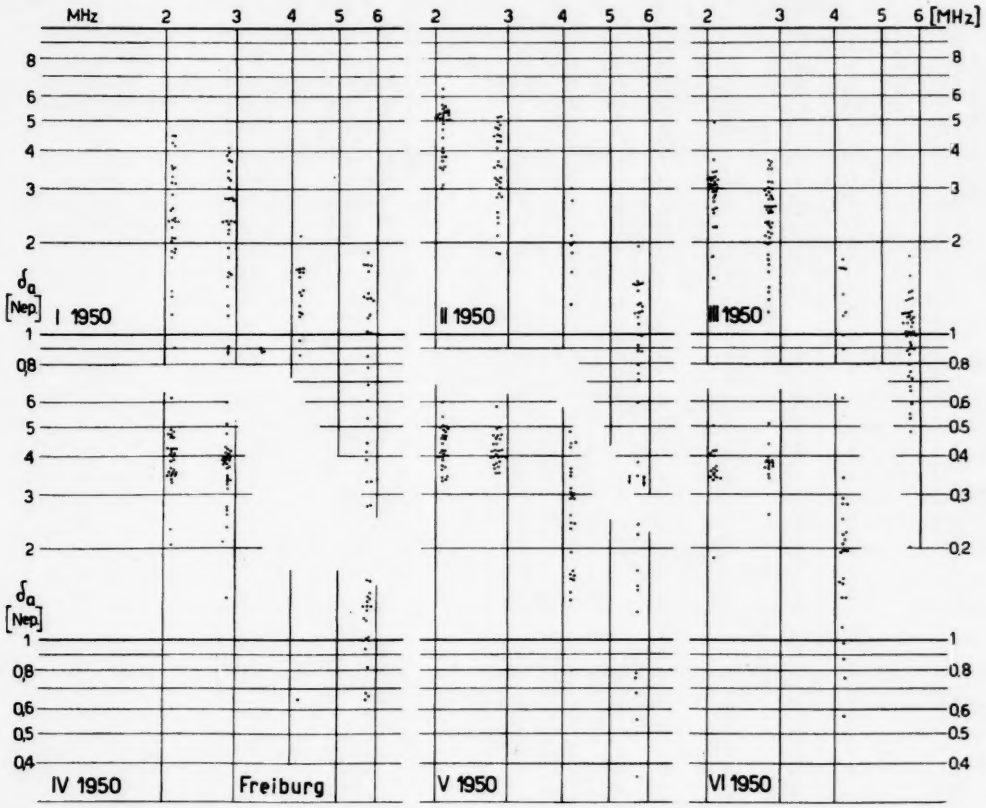


Fig. 2.

fonction de $\Delta_2 = (f + 1{,}2\ [\mathrm{MHz}])^{-2}$.
versus $\Delta_2 = (f + 1{,}2\,[\mathrm{MHz}])^{-2}$.

parcours du rayon à l'intérieur de la couche E.

$$\delta_3 = M \times \int k \times ds \qquad (M = \log e = 0{\cdot}4343). \tag{3}$$

Dans le cas de réflexion par la couche E, la limite supérieure de l'intégrale, à savoir la hauteur vraie de réflexion, varie avec la fréquence, selon la repartition de la densité électronique N_e.

D'autre part la variation du nombre de chocs avec l'altitude a aussi une influence importante.

Pour un certain nombre de modèles assez différents l'intégrale du décrement δ_3 se calcule exactement [3]. L'on constate que les différentes formes étudiées conduisent à des résultats peu différents. Une variation parabolique de la densité électronique avec variation exponentielle du nombre de chocs peut donc être adoptée comme s'écartant peu de la réalité. Ce cas peut être traité à l'aide de certaines

VOL. 2 1952

fonctions transcendantes [4], par contre nous avons ici effectué nos calculs numériques par développement selon le procédé indiqué par un des auteurs [5]. Nous avons d'ailleurs confronté les résultats numériques obtenues en suivant ce procédé avec ceux du calcul exact. La concordance est parfaite dans tout le domaine où les

Tableau 1. Décrement logarithmique de l'absorption, médiane mensuelle de Slough (Nepers)

Logarithmic absorption decrements, monthly medians of Slough (Nepers)

Mois	Fréquence [MHz]					
	2·0	2·4	2·8	3·2	4·0	4·8
1949						
I	6·0	4·5	4·0	3·5	2·4	2·2
II	5·1	3·7	3·6	4·0	2·35	2·15
III	4·5	3·6	3·1	3·4	2·4	1·9
IV	5·05	3·8	3·5	3·45	3·3	2·5
V	4·7	3·7	3·8	4·0	2·7	2·5
VI	4·8	4·3	4·5	5·0	3·6	3·6
VII	4·95	4·55	4·65	5·0	3·75	3·45
VIII	5·15	4·4	4·4	4·3	3·65	3·05
IX	4·3	3·9	4·1	4·1	3·15	2·9
X	4·4	4·0	3·95	3·75	2·7	2·25
XI	4·25	3·45	3·65	3·85	1·9	1·9
XII	4·1	3·6	4·1	3·3	2·0	1·8
1950						
I	3·9	3·2	3·7	3·0	1·75	1·4
II	4·7	4·3	4·3	4·7	2·3	2·0
III	4·05	3·4	3·6	4·6	2·4	1·9
IV	4·6	4·1	4·2	4·6	3·5	2·3
V	4·8	4·4	4·0	4·3	2·9	2·45
VI	4·3	4·3	4·05	4·4	3·35	2·4
VII	4·4	3·95	4·0	4·6	2·8	2·55
VIII	4·5	4·1	3·6	4·5	2·4	2·2
IX	4·35	4·55	4·1	4·6	2·6	2·25
X	3·6	4·0	3·95	4·3	2·0	1·4
XI	4·1	3·7	4·0	3·7	1·95	1·5
XII	5·1	5·0	4·0	4·0	1·9	1·7

Tableau 2. Décrement logarithmique de l'absorption, médiane mensuelle de Fribourg (logarithmes de BRIGG)

Logarithmic absorption decrements, monthly medians of Freiburg (BRIGGs logarithm)

Mois	Fréquence [MHz]			
	2·6	2·8	4·0	5·6
1948				
XI	1·45	1·08	(0·53)	0·45
XII	1·73	1·12	(0·70)	0·25
1949				
I	2·4	1·65	(0·55)	(0·40)
II	1·75	1·18	—	0·40
III	1·75	1·20	—	0·36
IV	1·90	1·30	0·96	0·60
V	1·85	1·32	0·98	0·58
VI	1·95	1·70	1·02	(0·60)
VII	2·50	1·98	0·96	—
VIII	2·30	1·53	0·75	(0·73)
IX	1·75	1·40	1·00	0·55
X	1·48	0·87	—	0·54
XI	1·25	1·00	0·53	0·40
XII	1·43	1·15	0·67	0·48
1950				
I	0·79	1·02	0·62	0·45
II	2·10	1·55	(0·88)	0·48
III	1·30	1·10	(0·60)	0·45
IV	1·77	1·68	—	0·53
V	1·80	1·80	1·20	0·55
VI	1·58	1·70	0·87	—
VII	1·80	1·80	0·92	0·52
VIII	2·10	1·82	1·12	(0·48)
IX	1·48	1·31	1·12	0·45
X	1·15	1·00	0·66	0·40
XI	1·55	1·73	0·70	0·30
XII	1·78	1·80	0·54	0·27

valeurs numériques des fonctions spéciales ont été évaluées [3]. Nous supposons donc une variation parabolique de la densité électronique (demi-épaisseur Y_m).

$$N_e = N_{eo} \times (1 - (Y/Y_m)^2) = Neo \times (1 - y^2). \qquad (4)$$

D'autre part, comme notre développement est indépendant de la repartition des chocs en fonction de l'altitude Y, on peut admettre une repartition très générale:

$$\nu = \nu_0 \times \psi\,(Y/Y_m) = \nu_0 \times \psi(y). \qquad (5)$$

Si la fonction ψ peut être développée en série suffisamment convergente dans l'intervalle considéré, l'intégration du produit $\nu \times N_e$ peut être effectuée terme pà terme.

Tableau 3. Méthode de calcul de l'absorption sélective
Calculation method of deviative absorption

$$\delta_3 = \frac{M}{2c} \int_{\substack{\text{aller et} \\ \text{retour}}} \frac{\nu}{\mu} \frac{f_0^2}{f^2}\, ds = \frac{M \times Y_m}{c} \nu_0 \int_{\text{aller}} \frac{\psi(y) f_0^2(y)}{f^2 \times \mu} \times dy$$

avec

$$y = Y/Y_m \,(s = Y) \quad \text{et} \quad \nu = \nu_0 \times \psi(y)$$

$$\delta_3 = \frac{M \times Y_m \times \nu_0}{c} \times \Delta_3 .$$

Introduisons:

$$f_0^2(y) = f_c^2 \times (1 - y^2) \quad \text{et posons} \quad \eta = f_c/f:$$

$$\Delta_3 = \eta^2 \int_{-1}^{+1 \text{ resp.} - \sqrt{1-\eta^{-2}}} \frac{(1-y^2) \times \psi(y) \times dy}{\sqrt{1 - \eta^2 (1-y^2)}} .$$

Cas de traversée: Cas de réflexion:

Posons: $\psi(y) = \sum_{i=0}^{\infty} a_i y^i$

$$\text{et } c_{it} = \eta^2 \int_{+1}^{0} \frac{y^i (1-y^2) \times dy}{\sqrt{1 - \eta^2 (1-y^2)}} \qquad \text{resp.} \qquad c_{ir} = \eta^2 \times \int_{+1}^{\sqrt{1-\eta^{-2}}} \frac{y^i \times (1-y^2) \times dy}{\sqrt{1 - \eta^2 \times (1-y^2)}}$$

nous obtenons ainsi:

$$\Delta_{3t} = 2 \times \sum_{i=0;\,2}^{\infty} a_i c_i \qquad \Delta_{3r} = \sum_{i=0}^{\infty} (-1)^i a_i c_i .$$

(i *pair*) (i toute valeur entière)

Le calcul est résumé dans la table 3. A l'aide de cette table l'évaluation peut être effectuée rapidement pour une fonction ψ quelconque.

Supposons maintenant pour ν une échelle de hauteur de 0,4 Y_m. Le décrement de l'absorption sélective, δ_3 s'écrit alors:

$$\delta_3(\eta) = \frac{M}{c} \times \nu_0 \times Y_m \times \Delta_3(\eta) = N \times \Delta_3(\eta) . \tag{6}$$

$$\eta = f_c/f .$$

Les tables 4 et 5 donnent les valeurs numériques de la fonction $\Delta_3(\eta)$.

3. Nouvelle Interprétation des Résultats Expérimentaux

Si l'absorption sélective n'est pas négligeable, le décrément d'absorption δ_a est égal à la somme de deux termes, dont le premier δ_2 correspond à l'absorption non-sélective causée par la couche D. Nous écrivons*

$$\delta_2 = B \times (f + f_L)^{-2} = B \times \Delta_2(f) . \tag{1a}$$

* Il nous semble préferable d'utiliser ici une lettre différente de A; ceci est appliquée dans le sens de l'équation (1) par accord international. D'ailleurs nous notons ici les décrements en logarithmes de Brigg et non en décibels.

VOL. 2 1952

Tableau 4. Décrement de l'absorption sélective calculé suivant le tableau 3, cas de la traversée
Decrement of deviative absorption, calculated with Table 3 for penetration

	c_0	c_2	c_4	c_6	c_8	c_{10}	Δ_3*
0·999	3·301	0·247	0·082	0·042	0·025	0·017	8·44
0·998	2·957	0·244	0·082	0·042	0·025	0·017	7·73
0·997	2·749	0·242	0·082	0·042	0·025	0·017	7·31
0·995	2·499	0·238	0·082	0·041	0·025	0·017	6·78
0·993	2·326	0·235	0·081	0·041	0·024	0·017	6·42
0·990	2·145	0·231	0·080	0·041	0·024	0·017	6·02
0·985	1·942	0·225	0·078	0·040	0·023	0·017	5·57
0·980	1·800	0·217	0·077	0·040	0·023	0·017	5·24
0·975	1·684	0·212	0·076	0·039	0·023	0·016	4·97
0·970	1·592	0·208	0·074	0·039	0·023	0·015	4·75
0·960	1·448	0·197	0·072	0·037	0·023	0·015	4·39
0·950	1·334	0·189	0·070	0·036	0·022	0·015	4·10
0·940	1·242	0·181	0·068	0·035	0·021	0·015	3·86
0·930	1·162	0·174	0·066	0·034	0·021	0·014	3·65
0·920	1·090	0·167	0·063	0·034	0·020	0·014	3·46
0·910	1·035	0·160	0·062	0·032	0·020	0·013	3·30
0·900	0·980	0·156	0·059	0·032	0·019	0·013	3·15
0·890	0·933	0·149	0·057	0·031	0·019	0·013	3·01
0·880	0·887	0·145	0·055	0·030	0·018	0·013	2·88
0·870	0·846	0·139	0·055	0·029	0·017	0·013	2·76
0·860	0·808	0·134	0·053	0·028	0·017	0·012	2·65
0·850	0·773	0·130	0·051	0·027	0·017	0·012	2·54
0·840	0·739	0·127	0·049	0·026	0·016	0·012	2·44
0·830	0·709	0·121	0·048	0·026	0·015	0·012	2·35
0·820	0·681	0·116	0·047	0·025	0·015	0·011	2·26
0·810	0·653	0·112	0·046	0·024	0·015	0·010	2·18
0·800	0·626	0·110	0·043	0·024	0·015	0·010	2·10
0·780	0·577	0·103	0·040	0·023	0·014	0·009	1·95
0·760	0·535	0·095	0·039	0·021	0·013	0·009	1·81
0·740	0·495	0·088	0·037	0·020	0·012	0·009	1·68
0·720	0·458	0·083	0·034	0·019	0·010	0·009	1·56
0·700	0·423	0·078	0·031	0·018	0·010	0·008	1·45
0·650	0·348	0·066	0·027	0·015	0·008	0·006	1·21
0·600	0·286	0·054	0·023	0·012	0·007	0·005	0·99
0·550	0·233	0·042	0·019	0·009	0·007	0·004	0·80
0·500	0·188	0·035	0·016	0·007	0·006	0·004	0·65
0·450	0·148	0·028	0·013	0·006	0·005	0·004	0·52
0·400	0·114	0·022	0·010	0·005	0·003	0·003	0·40
0·350	0·086	0·017	0·007	0·004	0·003	0·002	0·30
0·300	0·062	0·012	0·005	0·003	0·002	0·001	0·22
0·250	0·043	0·0085	0·0035	0·002	0·001	0·001	0·15
0·200	0·027	0·005	0·002	0·001	0·001	0·001	0·09
0·150	0·015	0·002	0·001	0·001	0·001	00·00	0·05
0·100	0·007	0·001	0·001	0·000	0·000	0·000	0·02
0·050	0·002	0·001	0·000	0·000	0·000	0·000	0·01

* Δ_3 dans le cas $\psi = \exp(-2{\cdot}5\,y)$, soit

$$\Delta_3 = 2\,c_0 + 6{\cdot}25\ c_2 + 3{\cdot}255\ c_4 + 0{\cdot}678\ c_6$$
$$+ 0{\cdot}0756\,c_8 + 0{\cdot}0052\ c_{10}.$$

Tableau 5. Décrement de l'absorption sélective calculé suivant le tableau 3, cas de réflexion
Decrement of deviative absorption, calculated with Table 3, for reflection

	c_0	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	Δ_3*
1·001	3·304	0·666	0·253	0·134	0·084	0·057	0·042	0·032	0·025	0·020	0·017	6·31
1·002	2·953	0·664	0·255	0·135	0·084	0·057	0·043	0·032	0·025	0·021	0·017	5·96
1·003	2·751	0·663	0·258	0·135	0·084	0·058	0·042	0·032	0·025	0·020	0·017	5·77
1·005	2·497	0·660	0·261	0·137	0·085	0·058	0·043	0·033	0·025	0·020	0·017	5·53
1·007	2·327	0·658	0·265	0·138	0·085	0·059	0·043	0·032	0·025	0·021	0·017	5·37
1·01	2·151	0·654	0·270	0·140	0·086	0·060	0·043	0·032	0·026	0·021	0·017	5·20
1·015	1·950	0·647	0·275	0·145	0·089	0·060	0·044	0·033	0·026	0·021	0·017	5·02
1·02	1·805	0·640	0·280	0·148	0·090	0·062	0·045	0·034	0·026	0·021	0·018	4·88
1·025	1·701	0·636	0·284	0·150	0·092	0·062	0·045	0·035	0·026	0·021	0·018	4·78
1·03	1·608	0·629	0·287	0·154	0·094	0·063	0·046	0·035	0·027	0·022	0·018	4·70
1·04	1·468	0·617	0·294	0·160	0·098	0·066	0·047	0·035	0·028	0·023	0·018	4·58
1·05	1·357	0·604	0·299	0·166	0·102	0·068	0·050	0·037	0·028	0·023	0·019	4·48
1·06	1·270	0·593	0·303	0·171	0·105	0·071	0·051	0·038	0·030	0·023	0·019	4·40
1·07	1·197	0·583	0·306	0·175	0·109	0·073	0·052	0·039	0·030	0·024	0·020	4·33
1·08	1·134	0·572	0·308	0·179	0·111	0·075	0·055	0·041	0·031	0·025	0·020	4·26
1·09	1·078	0·561	0·310	0·183	0·116	0·078	0·056	0·042	0·032	0·026	0·021	4·20
1·10	1·028	0·551	0·311	0·187	0·118	0·080	0·059	0·043	0·032	0·026	0·022	4·15
1·12	0·946	0·532	0·312	0·192	0·125	0·085	0·060	0·046	0·036	0·027	0·023	4·05
1·14	0·875	0·513	0·312	0·197	0·130	0·090	0·064	0·048	0·037	0·029	0·024	3·96
1·16	0·816	0·496	0·310	0·200	0·134	0·094	0·068	0·051	0·040	0·031	0·025	3·87
1·18	0·764	0·478	0·307	0·203	0·139	0·098	0·071	0·054	0·041	0·032	0·028	3·79
1·20	0·719	0·463	0·306	0·205	0·141	0·102	0·074	0·056	0·044	0·035	0·029	3·71
1·25	0·626	0·427	0·295	0·208	0·149	0·109	0·083	0·063	0·049	0·039	0·031	3·53
1·30	0·553	0·394	0·284	0·208	0·154	0·115	0·088	0·069	0·054	0·043	0·036	3·36
1·35	0·496	0·366	0·272	0·205	0·155	0·120	0·094	0·074	0·059	0·048	0·039	3·19
1·40	0·447	0·340	0·260	0·201	0·157	0·123	0·097	0·079	0·064	0·052	0·043	3·04
1·45	0·406	0·318	0·248	0·196	0·156	0·125	0·101	0·082	0·067	0·055	0·047	2·89
1·50	0·372	0·297	0·237	0·191	0·154	0·125	0·103	0·085	0·069	0·058	0·050	2·76
1·60	0·316	0·261	0·216	0·179	0·149	0·125	0·105	0·089	0·075	0·063	0·055	2·61
1·70	0·273	0·231	0·196	0·167	0·132	0·121	0·105	0·091	0·077	0·067	0·060	2·28
1·80	0·238	0·206	0·179	0·155	0·134	0·118	0·103	0·090	0·078	0·069	0·062	2·08
1·9	0·210	0·185	0·163	0·144	0·126	0·112	0·100	0·089	0·079	0·070	0·063	1·90
2·0	0·186	0·167	0·149	0·133	0·119	0·107	0·097	0·086	0·078	0·070	0·064	1·74
2·2	0·151	0·138	0·126	0·115	0·104	0·096	0·088	0·081	0·074	0·068	0·063	1·48
2·4	0·124	0·116	0·097	0·100	0·096	0·086	0·082	0·075	0·071	0·064	0·061	1·27
2·6	0·105	0·099	0·093	0·087	0·081	0·077	0·072	0·068	0·063	0·060	0·058	1·10
2·8	0·090	0·085	0·080	0·076	0·073	0·068	0·064	0·062	0·059	0·055	0·055	0·96
3·0	0·078	0·074	0·071	0·068	0·064	0·061	0·059	0·057	0·053	0·051	0·051	0·85
3·2	0·068	0·066	0·063	0·060	0·057	0·055	0·053	0·051	0·049	0·047	0·045	0·75
3·4	0·060	0·058	0·055	0·053	0·051	0·050	0·048	0·047	0·045	0·043	0·042	0·67
3·6	0·054	0·052	0·050	0·048	0·047	0·045	0·043	0·043	0·040	0·040	0·039	0·60
3·8	0·048	0·046	0·045	0·043	0·041	0·041	0·040	0·039	0·038	0·037	0·036	0·54
4·0	0·043	0·042	0·040	0·039	0·037	0·037	0·036	0·035	0·035	0·034	0·034	0·48
4·5	0·033	0·032	0·031	0·030	0·031	0·030	0·029	0·029	0·028	0·040	0·027	0·38
5·0	0·027	0·027	0·026	0·026	0·025	0·025	0·025	0·025	0·024	0·023	0·023	0·31
5·5	0·023	0·022	0·022	0·021	0·021	0·021	0·021	0·020	0·020	0·020	0·020	0·26

* Δ_3 dans le cas $y = \exp(-2{\cdot}5y)$, soit

$$\Delta_3 = c_0 + 2{\cdot}5\ c_1 + 3{\cdot}125\ c_2 + 2{\cdot}604\ c_3 + 1{\cdot}628\ c_4 + 0{\cdot}814\ c_5 + 0{\cdot}339\ c_6 + 0{\cdot}121\ c_7 + 0{\cdot}038\ c_8 + 0{\cdot}0105\ c_9 + 0{\cdot}0026\ c_{10}.$$

VOL. 2 1952

Tableau 5 (Suite)

	c_0	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	Δ_3*
6·0	0·019	0·019	0·019	0·018	0·018	0·018	0·017	0·018	0·017	0·017	0·017	0·22
6·5	0·016	0·016	0·016	0·016	0·015	0·015	0·015	0·015	0·015	0·015	0·015	0·19
7·0	0·014	0·014	0·014	0·013	0·013	0·013	0·012	0·013	0·012	0·012	0·012	0·16
7·5	0·012	0·012	0·012	0·011	0·011	0·011	0·011	0·011	0·011	0·011	0·011	0·14
8·0	0·011	0·011	0·011	0·011	0·010	0·011	0·010	0·010	0·010	0·010	0·010	0·13
9·0	0·008	0·008	0·008	0·008	0·008	0·008	0·008	0·008	0·008	0·008	0·008	0·10
10·0	0·006	0·007	0·007	0·007	0·007	0·007	0·007	0·007	0·007	0·007	0·007	0·08
15·0	0·003	0·003	0·003	0·003	0·003	0·003	0·003	0·003	0·003	0·003	0·003	0·03
20·0	0·002	0·002	0·001	0·001	0·001	0·001	0·001	0·001	0·001	0 001	0·001	0·02
30·0	0·001	0·001	0·001	0·001	0·001	0·001	0·001	0·001	0·001	0·001	0·001	0·01
40·0	0·000	0·001	0·001	0·001	0·001	0·001	0·000	0·000	0·000	0·000	0·000	0·00
50·0	0·000	0·001	0·000	0·000	0·000	0·000	0·000	0·000	0·000	0·000	0·000	0·00

* Δ_3 dans le cas $\psi = \exp(-2{\cdot}5y)$, soit

$$\Delta_3 = c_0 + 2{\cdot}5\,c_1 + 3{\cdot}1\,25c_2 + 2{\cdot}604\,c_3 + 1{\cdot}628\,c_4 + 0{\cdot}814\,c_5 + 0{\cdot}339\,c_6 + 0{\cdot}121\,c_7 + 0{\cdot}038\,c_8 + 0{\cdot}0105\,c_9 + 0{\cdot}0026\,c_{10}.$$

Le deuxième terme, δ_3, est l'absorption sélective dans la couche E (voir l'équation[6]).

$$\delta_a(f) = B \times \Delta_2(f) + N \times \Delta_3(\eta). \tag{7}$$

Les mesures sont faites sur plusieures fréquences; nous pouvons calculer les caractéristiques de l'absorption, B et N, en utilisant les mesures faites sur deux quelquonques des fréquences (f est connue, $f_L = 1{,}2$ MHz en Europe; $f_c = f_0 E$ peut être trouvé dans les médianes mensuelles publiées.)

Il nous a semblé commode d'utiliser un procédé graphique pour la détermination des caractéristiques N et B. Nous tracons la droite représentant l'équation (8) en axes N et B. Cette droite coupe les axes aux points,

$$\delta_a/\Delta_3 \quad \text{et} \quad \delta_a/\Delta_2.$$

Les mesures déterminent ainsi une droite par fréquence. Toutes ces droites devraient se couper en un même point. Les coordonnées de ce point indiqueraient les valeurs effectives de B et N, soit de l'absorption non-sélective et sélective.

Nous avons dépouillé de cette facon les mesures de Slough et de Fribourg obtenues en 1949 et 1950. Les graphiques dont les Fig. 3 et 4 donnent de exemples n'ont pour base que les médianes mensuelles des valeurs δ_a des différentes fréquences. Dans la plupart des cas on obtient plusieurs points d'intersection. Il existe des mois, surtout d'hiver où l'intersection ne correspond guère à notre théorie. Peut-être l'existance de la couche E_2 a-t-elle une certaine influence ? Par ailleurs l'existence d'une couche E_s rend les mesures difficiles, soit par réflexion partielle, soit par annulation de l'absorption selective dans le cas d'un rayon réflechi au dessous du centre de la couche E par une couche E_s. Parfois même des erreurs d'interprétation peuvent avoir eu lieu. Il est intéressant que les résultats aberrants se trouvent surtout sur les fréquences qui sont gênées par les influences de E_2 et E_s

(4:8 MHz à Slough, 4:0 à Fribourg).

Considérons plus es détail ces influences. Une réflection sur la couche E_2 peut facilement être confondue avec l'écho recu après deux réflexions successives sur la couche E_s. Dans ces conditions la réflexion partielle sur la couche E_s est interprétée par erreur comme un effet d'absorption. Une réfléxion ordinaire sur la couche *F* simultanée d'une réflexion extraordinaire venant de la couche *E* peut aussi être interprété par erreur comme écho ayant subi deux réflexions sur la couche E_s. Dans ces cas on obtiendrait une absorption trop faible. Enfin la méthode de dépouillement indiquée ci dessus suppose l'existence d'une absorption sélective pour les fréquences supérieures à la fréquence critique de la couche *E*. Or s'il existe une couche E_s dont l'ionisation est telle qu'elle occulte complètement la couche *F* pour la fréquence considérée, on obtient une valeur correcte du coefficient d'absorption, mais le parcours n'entre jamais dans la région centrale de la couche *E*. En conséquence on dévrait interpréter un résultat obtenu dans ces conditions seulement en termes d'absorption δ_2 introduite par le passage à travers la couche *D*. Ici l'intertroduction de δ_3 n'est donc pas justifiée.

Afin d'éviter ces particularités nous avons cherché dans chacun des diagrammes mensuels la médiane de tous les points d'intersection (à condition que les valeurs correspondantes de *B* et *N* soient positives). Cette

Tableau 6. Valeurs mensuelles de B et N obtenues à l'aide des diagrammes d'intersection.
Monthly values of B and N as obtained from the intersection diagrams

		Mois												médianes
		I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*	*IX*	*X*	*XI*	*XII*	
B 1948	Sl.	—	—	—	—	—	—	—	—	—	—	—	—	—
	Fr.	—	—	—	—	—	—	—	—	—	—	(5·5)	(16·5)	—
1949	Sl.	21·5	13·5	15·5	16·0	15·0	15·0	15·0	17·5	13·5	12·0	12·5	13·5	15·0
	Fr.	(16·5)	(17·0)	—	(15·5)	17·0	21·5	(23·5)	(21·0)	14·5	—	8·0	12·0	16·7
1950	Sl.	12·5	15·0	9·0	12·5	17·0	12·0	12·0	16·0	13·0	12·5	12·0	16·0	12·5
	Fr.	—	21·0	10·5	12·5	12·0	10·5	13·0	15·5	11·5	6·0	(4·5)	(11·5)	11·5
													médiane	13·8
N 1948	Sl.	—	—	—	—	—	—	—	—	—	—	—	—	—
	Fr.	—	—	—	—	—	—	—	—	—	—	(0·165)	(0·050)	—
1949	Sl.	0·115	0·210	0·155	0·185	0·270	0·315	0·355	0·255	0·285	0·240	0·215	0·260	0·247
	Fr.	(0·135)	(0·035)	—	(0·100)	0·105	0·070	(0·095)	(0·140)	0·165	—	0·110	0·085	0·102
1950	Sl.	0·180	0·250	0·305	0·340	0·220	0·325	0·280	0·180	0·280	0·235	0·185	0·205	0·243
	Fr.	—	0·075	0·130	0·250	0·305	0·275	0·265	0·255	0·165	0·180	(0·425)	(0·310)	0·255
													médiane	0·245

VOL. 2 1952

médiane (point central des points d'intersection) a été indiquée par une petite croix. Les valeurs de B et N ainsi obtenues sont tabulées dans la table 6.

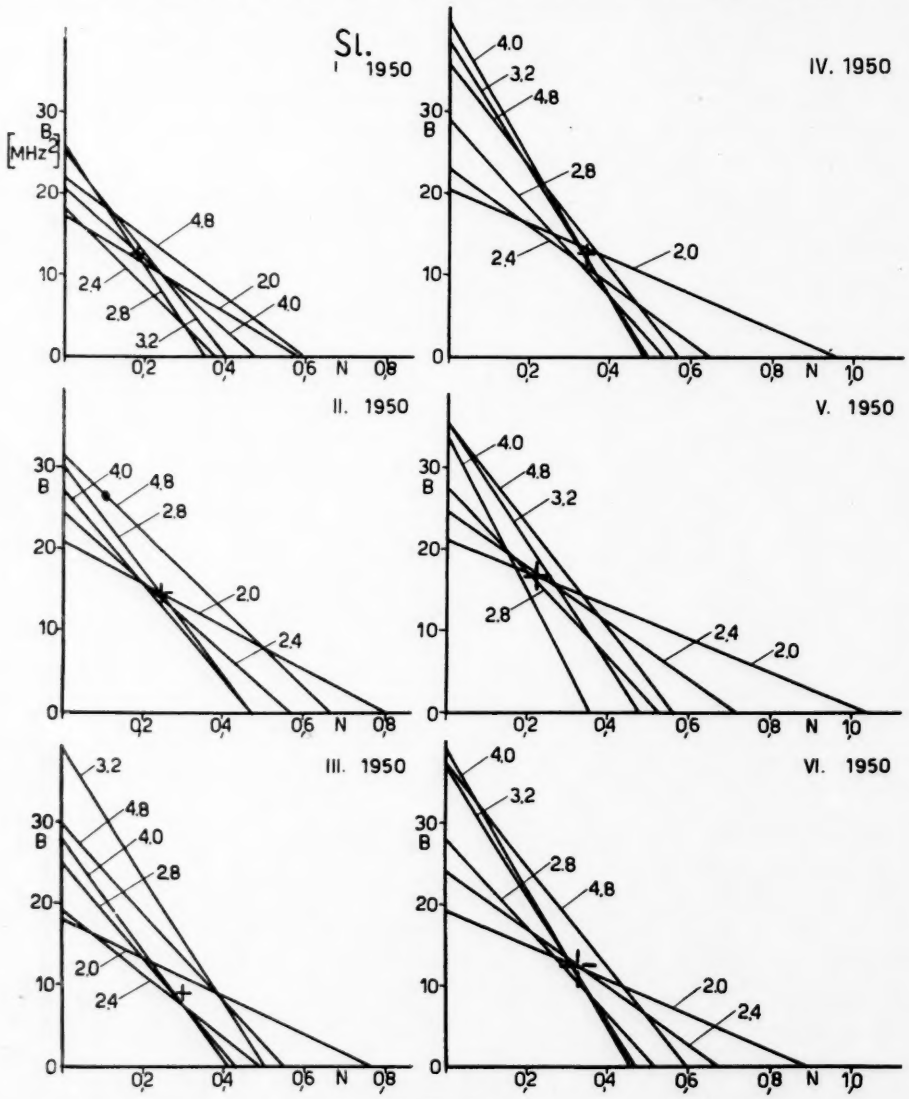


Fig. 3.

Fig. 3 et 4. Diagrammes d'intersection (B caractérisant
Intersection diagrams, each line corresponding to a certain frequency

A titre de contrôle nous avons répété notre dépouillement à l'aide des moyennes annuelles des valeurs mensuelles δ_a/Δ_3 et δ_a/Δ_2*. Ainsi nous obtenons

* Nous avons choisi ce procédé pour éliminer les irrégularités trouvées pendant certains mois sur certaines fréquences.

la Fig. 5. A part les fréquences déjà mentionnées (4,0 resp. 4,8 MHz), l'intersection est assez satisfaisante. Les résultats de Fribourg de l'année 1949 semblent être moins cohérents qu'en 1950*. La comparaison des valeurs de N obtenues ainsi

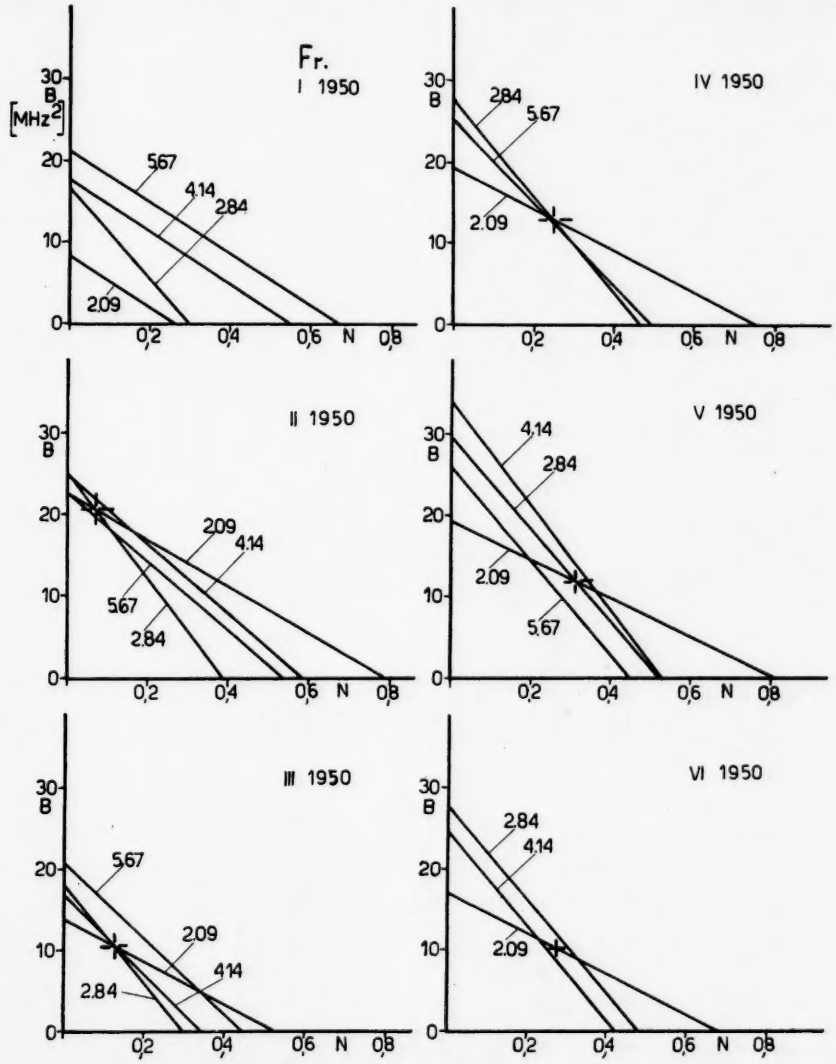


Fig. 4.

l'absorption non-selective, N l'absorption sélective).
(B and N characterise non-deviative and deviative absorption).

(d'après la Fig. 5) avec les valeurs mensuelles de la table 6 confirme cette impression. Toutes les valeurs obtenues pour N ont été notées dans la Fig. 6. A part

* En fait les méthodes de mesure et de dépouillement ont été changées plusieures fois au cours de cette année. La qualité de l'observation était certainement moins bonne pendant l'année 1949.

VOL. 2 1952

les mesures de Fribourg en 1949 la valeur annuelle de N (obtenue à l'aide de la Fig. 5) est pratiquement identique à la moyenne des valeurs mensuelles correspondantes, ce qui nous semble être satisfaisant.

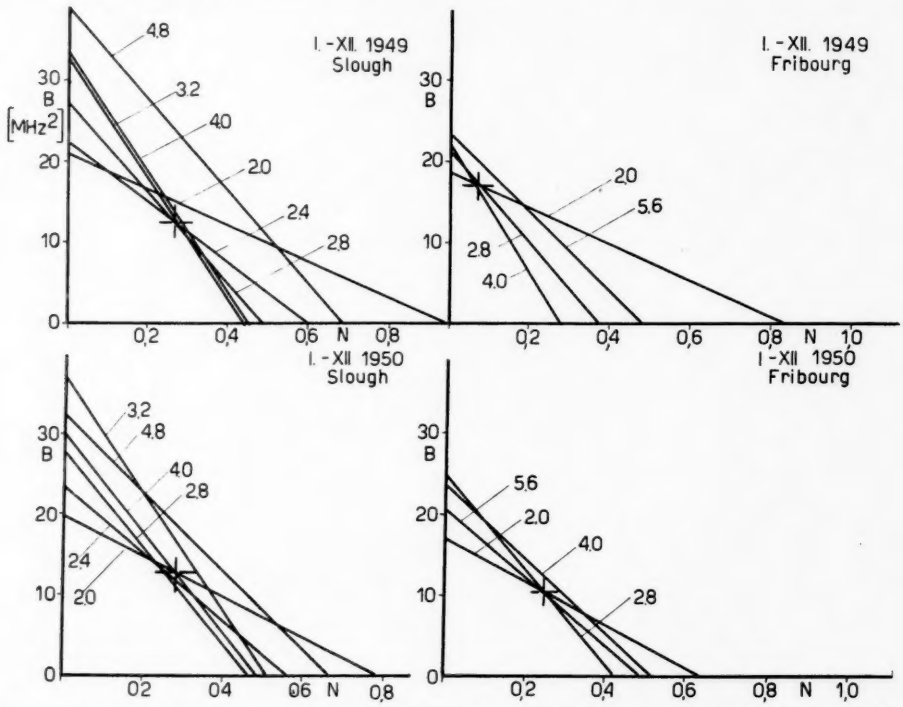


Fig. 5. Diagrammes d'intersection obtenues à l'aide des médianes annuelles des δ_a/Δ_2 resp. δ_a/Δ_3.
Intersection diagrams obtained with yearly medians of the δ_a/Δ_2 resp. δ_a/Δ_3.

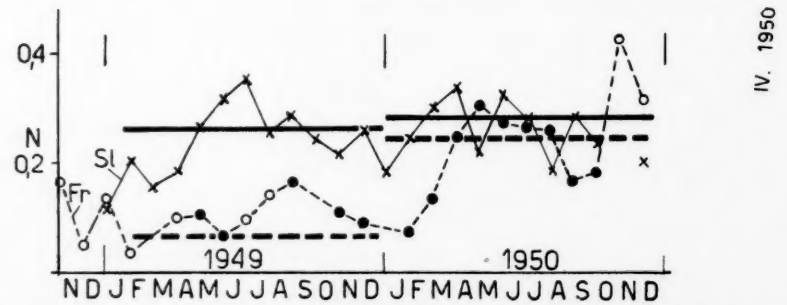


Fig. 6. Valeur N caractérisant l'absorption sélective de la couche E. (Traits épais: valeurs obtenues à l'aide de la Fig. 5.)
Deviative absorption occurring in the E-layer. (The horizontal lines indicate values obtained from Fig. 5).

Enfin nous avons la Fig. 7 indiquant la variation des valeurs B. La Fig. 8 est le corrélogramme correspondant. Nous trouvons une corrélation assez bonne (le „nombre" de corrélation [1] est 0,88). Ici la différence des résultats de Fribourg

et de Slough est moins importante. Les médianes des B mensuel sont sensiblement indentiques, soit 13,5 pour Slough et 13 pour Fribourg*. La Fig. 7 révèle une

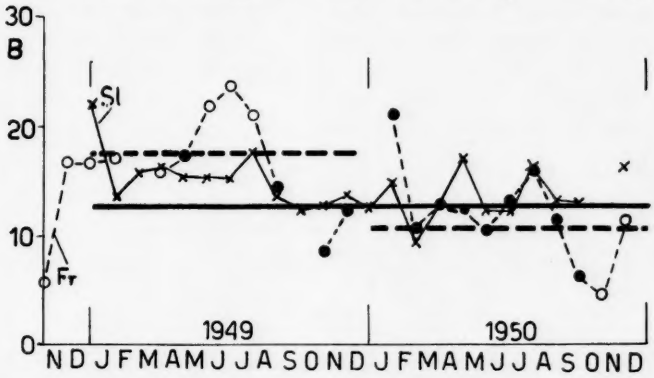


Fig. 7. Valeur B caractérisant l'absorption non-sélective de la région D. (Traits épais: valeurs obtenues à l'aide de la Fig. 5.)
Non-deviative absorption of the D-layer. (The horizontal lines indicate values obtained from Fig. 5.)

influence saisonnière caractérisée par un minimum en novembre (en accord avec nos conclusions antérieures tirées des mesures de Slough 1935—1947 [6]. La diminution de l'absorption avec la décroissance du cycle solaire est très remarquable.

VOL. 2 1952

4. Propositions Pour L'Évaluation Future des Mesures D'Absorption

Il résulte de ce qui précède que les valeurs de l'absorption attribuées en général à la couche D sont dues partiellement à l'absorption sélective de la couche E. La méthode de dépouillement habituelle néglige cette dernière influence; en conséquence on attribue à la couche D une trop forte absorption.

Nous proposons de distinguer, dans l'évaluation, les influences des couches D et E.

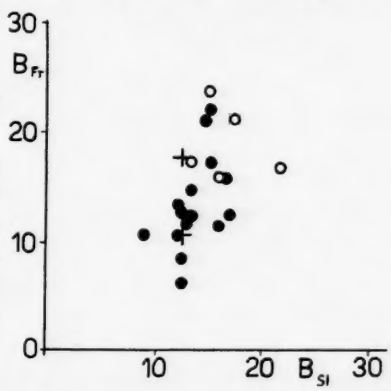


Fig. 8. Corrélogramme des valeurs B obtenues à Slough et à Fribourg. (Les croix indiquent les valeurs annuelles obtenues à l'aide de la Fig. 5.)
Correlogram of B-values obtained at Slough and at Freiburg. (The crosses indicate yearly values obtained from Fig. 5.)

Ceci serait possible par le procéde graphique que nous venons d'indiquer (Fig. 3 à 5): La valeur du décrement δ_a observée sur une certaine fréquence devrait être divisée par $\Delta_2(f)$ resp. $\Delta_3(f/f_c)$ afin d'obtenir les points d'intersection sur les axes B et N. De cette facon une droite par fréquence serait définie. Or d'après notre experience il est assez invraisemblable pour une mesure isolée que les droites s'acommodent avec un point d'intersection plus ou moins défini. Nous n'avons même pas eu des résultats entièrement satisfaisants avec les médianes mensuelles des δ_a.

Nous croyons qu'un autre procédé s'adapte mieux à l'insuffisance des mésures. La Fig. 6 nous a appris que la valeur N correspondant à l'influence de la couche

* Ce qui correspond à 270 resp. 260 [décibels $\times$ MHz^2].

E ne varie guère dans le temps. Si nous utilisons la valeur de 0,25 obtenue avec les mesures discutées dans le paragraphe 3 nous ne risquons pas trop de nous tromper. Introduisant cette valeur dans le diagramme indiqué cidessus nous trouvons par intersection avec les droites une série de valeurs B correspondant au différentes fréquences et dont la médiane peut être prise comme valeur effective de B. Il est probable que cette méthode éliminerait une partie des difficultés qu'on trouve actuellement dans l'évaluation des données d'absorption.

Les irrégularités qui subsisteraient seraient vraisemblablement à attribuer à des effects d'optique géométrique (focalisation et défocalisation). Celles-ci ne peuvent guère être éliminées au cours d'une mesure isolée [1]. Ceci semble être la raison pour laquelle les résultats de telles mesures ne sont pas toujours cohérents.

5. Le Nombre de Choc Dans la Couche E

Nous avons trouvé dans le paragraph 3 qu'il résulte des mesures de l'absorption effectuées à Slough et à Fribourg une valeur de N égale 0,25 or nous avons

$$N = \frac{M}{c} \times \nu_0 \times Y_m , \tag{8}$$

Y_m la demi-épaisseur de la couche E est comprise entre 25 et 30 km. En conséquence le nombre de chocs valable pour cette couche, ν_0, peut être

$$\underline{\nu_0 = 6 \ldots 7 \times 10^3 [s^{-1}]}.$$

Cette valeur correspond au centre de la couche E, soit à une altitude d'à peu près 125 km.

Il est très probable qu'à cette altitude seuls les chocs avec les molécules neutres sont importants [7]. Dans ce cas nous pouvons calculer ν d'après la formule cinétique:

$$\nu = \left(\frac{\pi k \times D^4}{2 m_e}\right)^{\frac{1}{2}} \times N_m \times \sqrt{T} = 4{,}7 \times 10^{-16} [m^3 s^{-1} \, \mathrm{grd}^{-\frac{1}{2}}] \times N_m \times \sqrt{T} . \tag{9}$$

N_m la densité moléculaire; T la température absolue; nous avons introduit une valeur de la section efficace des molécules de $7{,}5 \times 10^{-20}$ [m^2].

Si nous adoptons pour l'altitude de 125 km comme température la valeur de 370° K, qui est conforme aux mesures des V_2 [8], nous avons pour cette altitude:

$$N_m = \frac{6 \ldots 7{,}2 \times 10^3}{4{,}7 \times 10^{-16}} \times \frac{1}{\sqrt{370}} = 6{,}5 \ldots 8 \times 10^{17} [m^{-3}].$$

Cette valeur, correspondant à une valeur de la pression de $3 \ldots 4 \times 10^{-5}$ [mb], est en exellent accord avec les mesures des V_2 [8], qui ont donné $1, 3 \ldots 4 \times 10^{-5}$ [mb] pour l'altitude de 125 km. (Notons encore que le centre de la couche E semble s'être trouvé à une altitude inférieure (120 km) pendant le dernier minimum du cycle solaire).

Notre valeur du nombre de chocs est inférieure à la plupart des estimations antérieures obtenues a l'aide des mesures ionosdhériques. Nous croyons que notre

résultat (qui est basé sur un très grand nombre de mesures correspondant à environ 10^5 lectures d'amplitude), ait une certaine supériorité aux estimations antérieures qui n'ont été basées en général que sur des observations isolées. La valeur de la densité déduite de notre valeur du nombre de chocs est maintenant en bon accord avec les expériences modernes.

Bibliographie

[1] Rawer, K.; J. Atmosph. Terr. Phys. 1951 **2** 38. [2] Booker, H. G.; Proc. Roy. Soc. A, 1935 **150** 267–286. [3] Argence, E., Mayot, M. et Rawer, K.; Annales de Géophysique 1950 **6** 242–285. [4] Argence, E. et Rawer, K.; Compt. Rend. (Paris) 1949 **229** 996. [5] Bibl, K.; La stratification FO et les couches intermédaires entre les régions *E* et *F* de l'ionosphère, rapport SPIM—R 9, Service de Prévision Ionosphérique Militaire (1950). [6] Rawer, K.; Les paramètres de l'absorption des ondes par la couche *D*, et leur prévision, rapport SPIM—R 5, Service de Prévision Ionosphérique Marine (1949). [7] Majumdar, R. C.; Z. Phys. 1937 **107** 599–622. [8] Best, N. R., Havens, R. J. et Lagow, H.; Phys. Rev. 1947 **70** 915.

VOL.
2
1952

Journal of Atmospheric and Terrestrial Physics, 1951, Vol. 2. pp. 66 to 78. Pergamon Press Ltd., London

Ionospheric behaviour in the F_2 region at Singapore

B. W. Osborne, M. Sc.

Radio Research Station, Slough, Bucks

(Received 3 July 1951)

ABSTRACT

Previous work on ionospheric behaviour at low latitudes has demonstrated the existence of ionospheric phenomena peculiar to equatorial regions, particularly during the mid-day period and after sunset, within the region of a "trough" of low F_2 layer critical frequency near the magnetic equator. Observations at Singapore (on the southern edge of this "trough") from November 1948 have shown the inter-dependence of F_2 layer critical frequency and height with the seasonal occurrence of thick layer effects. Ionospheric behaviour at Singapore during the December solstice shows marked differences from that at the equatorial station of Huancayo, Peru, during daylight hours. Layer height changes after sunset, and the associated occurrence of the equatorial scatter phenomenon, are discussed. Consideration is given to the interpretation of vertical incidence virtual height measurements under Singapore conditions of a thick F region.

1. Introduction

The Japanese Physical Institute for Radio Waves established a chain of ionospheric observing stations in the South-East Asian region under Japanese domination during the period from 1942 to 1945, those outside Japan being under the administration of the Japanese Army or Navy. This chain included a Japanese Army station which was set up at Singapore in 1943, and though the original records are not now available, much of the information was later collated in Japan by the United States Army and was circulated in a report on Japanese research [1]. The Japanese measurements at Singapore were made with a manually-operated apparatus, which was used to obtain virtual height readings at predetermined frequency intervals. The F region at the Singapore station, during this period of minimum sunspot activity, showed frequent stratification, various virtual height levels being regularly obtained between F_1 and the highest F_2 layer. The problem of fitting a continuous $h'f$ curve over a series of manually-obtained points is not, however, always simple when the layer shape is irregular, and the interpretation given at any time must rest to some extent on the original observer.

Night scatter was observed at all the Japanese equatorial stations, including Singapore, this scatter, which occurred most frequently at the equinoctial periods, being similar in form to that previously observed at Huancayo [2], [3].

Some further experimental results obtained on the Japanese equipment have been given by Hon Yung Sen [4].

The Department of Scientific and Industrial Research established an ionospheric station at Singapore during the latter part of 1948, in the Raffles College grounds of the present University of Malaya, at the location 1° 19′ N, 103° 49′ E. The ionosphere was sounded at vertical incidence by the pulse-echo method of Breit and Tuve. Preliminary manual measurements commenced in October 1948, and an automatic photographic frequency-sweep recorder was installed and used from the first week of January 1949. This recorder, originally built to the design

of THOMAS and CHALMERS [5] was modified to follow the system originally due to GILLILAND [6]. In this system the transmitted frequency is obtained by beating the receiver local oscillator with a stable fixed-frequency pulsed oscillator operating at the intermediate frequency of the receiver, thus ensuring that the receiver is tuned to the frequency of the transmitted signal. The frequency band covered was from 2·2 Mc/sec to 16 Mc/sec, and the power output of the equipment, which varied over the band, did not exceed 200 watts.

Parallel vertical rhombic aerials were used for transmission and reception. The aerial system was erected in the magnetic NE—SW direction, so as not to discriminate between the ordinary and extraordinary components of the returned echoes, which are plane polarized parallel to the magnetic N—S and E—W directions respectively in equatorial regions.

Measurements of frequencies and virtual heights were obtained from the film records by standard methods, and, whenever possible, the values of the height of maximum ionization (*hm*) and the semi-thickness (*ym*) of the equivalent parabolic layer were derived by the method of APPLETON and BEYNON [7], [8].

2. IONOSPHERIC STRUCTURE OBSERVED DURING DAYLIGHT HOURS

Preliminary $h'f$ curves taken at Singapore frequently showed that the F region was not simple, and that very great delays were likely to be observed occasionally. Simple structure was found only between midnight and 0800 L.M.T., or in the daytime at some equinoctial periods during sunspot maximum years.

VOL. 2 1952

The minimum apparent virtual height of the upper F_2 shelf was frequently about 800 km, and sometimes first order reflections could be obtained from above 1400 km. It was therefore necessary to be able to record virtual heights up to at least 1200 km, though this meant some loss of accuracy in height measurement.

Three main types of F_2 behaviour could be identified, which were dominant at different periods of the year:

(a) the morning thick layer type,

(b) the afternoon thick layer type, and

(c) the equinoctial layer type.

2.1. *The morning thick layer phenomena*

From November to February large F_2 region group retardation delays were observed during the first two hours after dawn. The minimum height of the layer was normal and its shape could be represented by a parabola with reasonable accuracy but the semi-thickness of the parabola of best fit was exceptionally great, often of the order of 300 km. Thus the layer appeared to grow abnormally thick after dawn, the maximum ionization density being lower than would have been expected from comparable measurements at higher latitudes.

The separation of the ordinary and extraordinary critical frequencies, $f_0 F_2$ and $f_x F_2$ respectively, was abnormally large during this period, values of 0·7 Mc/sec being obtained for this single thick layer, instead of the normal 0·5 Mc/sec.

After about 0900 L.M.T., this thick layer became unstable and at times turbulent, with ledges in the F_2 region appearing on frequencies between 6 Mc/sec and 9 Mc/sec

at various heights up to 800 km. The apparent height "*hm*" of the maximum ionization density for the parabola of best fit, ignoring the effects of group retardation in the lower ledges, at times exceeded 800 km.

Although irregular ledges were often seen at frequencies below 7 Mc/sec, no regular F_1 layer could be identified during this period. The ledges were not well defined, and the F_1 and F_2 layer ionizations overlapped extensively, thus preventing the measurement of the actual critical frequencies of F_1 and the ledges. The F_1 layer did, however, become prominent during periods of abnormally low ionospheric ionization. The phenomena disappeared after noon with the establishment of a stable type of layer which is described in detail in Section 2.2 below.

2.2 *The afternoon thick layer phenomena*

A stable layer structure, which more closely resembled the structure obtained during summer months at higher latitudes, was normal in the afternoon period during the months June to August. During this period, comparatively regular ledge formation similar to that associated with the F_1 layer could be observed occasionally. The F_2 layer was considerably thicker than is normal at higher latitudes, semi-thicknesses of the order of 300 km being found at times, but the height of the maximum, "*hm*," remained about 500 km or less, in contrast with the very high values found with the morning phenomena (2.1). The turbulent conditions and subsidiary ledges found in the morning were usually conspicuously absent.

As is usual at higher latitudes in the summer, the critical frequency of the layer was comparatively low during these conditions. This type of variation is frequently observed at Huancayo [2], [3].

2.3 *Equinoctial layer type*

At equinoctial periods, from March to May and from September to October, the F region ionization generally formed a single layer throughout the day. A measurable F_1 critical frequency was seldom observed, although the F_2 structure was usually sufficiently simple to permit observation of any distinct F_1 region.

Data obtained between 1948 and 1951 indicated that the thickness of the layer tends to increase towards sunspot minimum.

The critical frequency of this single F layer was usually considerably greater than the critical frequencies found at similar times when the other types of structure were present, though they are lower than would be expected from the trends at higher latitudes.

3. Seasonal Variations of the F_2 Region Ionization over Singapore

As indicated in Figure 1a, the monthly mean noon values of the critical frequency of the F_2 layer over Singapore show both annual and semi-annual variations, unequal maxima appearing during the equinoctial months of April and October when the sun is near the zenith.

The corresponding values of the height of maximum ionization, *hm*, Figure 1b, show a dominant annual variation, the maximum values of *hm* being obtained during the months December to February. This indicates that Singapore must be considered to be a southern hemisphere station, as might be expected where the magnetic dip is 15° S.

Both of these parameters, f_0F_2 and hm, are sensitive to the effects of the thick layer phenomena described above, which modify, therefore, the apparent annual and diurnal variation of the layer constants. It is convenient to construct a thick layer frequency occurrence parameter, T, which will indicate the probability of the formation of a thick layer at any given time. For this purpose an arbitrary definition has been established which states that thick layer phenomena are present when either multiple ledges are observed in the F region, or the virtual height with frequency curve shows gradients exceeding 150 km per Mc/sec at frequencies appreciably lower than those affected by the severe group retardation near the maximum critical frequency. The ratio, T, of the total number of these occurrences to the total number of useful observations at the particular local mean time in any month then indicates the frequency of occurrence of thick layer phenomena. A curve giving the monthly variation of the parameter T at noon, expressed as a percentage probability of occurrence, is shown in Figure 1c. Figure 1 indicates that minimum values of f_0F_2 occur when T is a maximum, and that T and f_0F_2 are inversely associated.

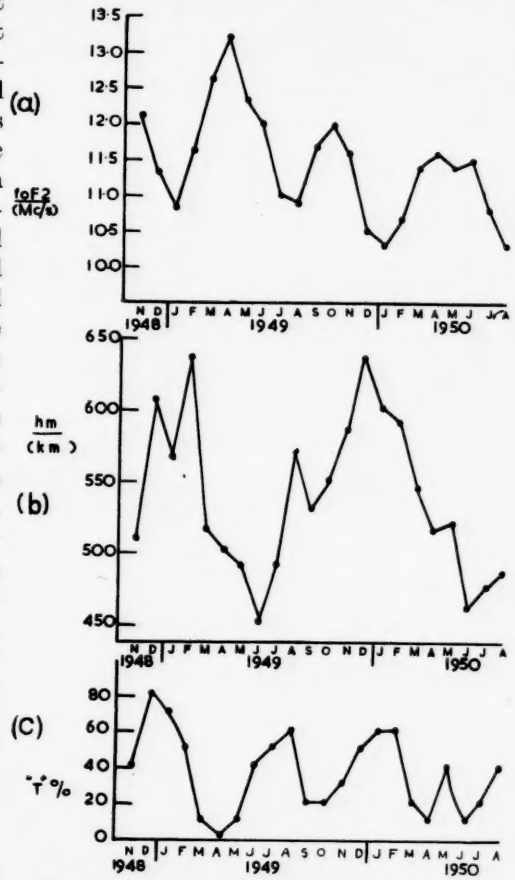


Fig. 1. Seasonal variations at Singapore at noon.

The diurnal variation of thick layer occurrence T for each month of the year 1949 is shown in Figure 2, which also illustrates the diurnal and seasonal variation of the phenomena discussed in Sections 2.1 and 2.2 above. It can be seen that thick layer effects are confined to daylight hours, with minima at the equinoctial and maxima at the solsticial periods. The marked increase due to the morning thick layer phenomena is evident between the hours of 0800 and 1100 L.M.T. from October to January, giving a value of T as high as 80% at 1000 L.M.T. in December 1949.

The monthly mean values of f_0F_2 at Singapore from November 1948 to January 1950 inclusive are shown in Figure 3 in the form of a contour map. From this diagram, it is possible to obtain either the observed diurnal variation for a given month, or the seasonal variation at a given local mean time, by taking vertical and horizontal cross-sections respectively. Cyclic variations of the monthly mean critical frequency, with a period of six months, are clearly evident at nearly all hours of the day.

VOL. 2 1952

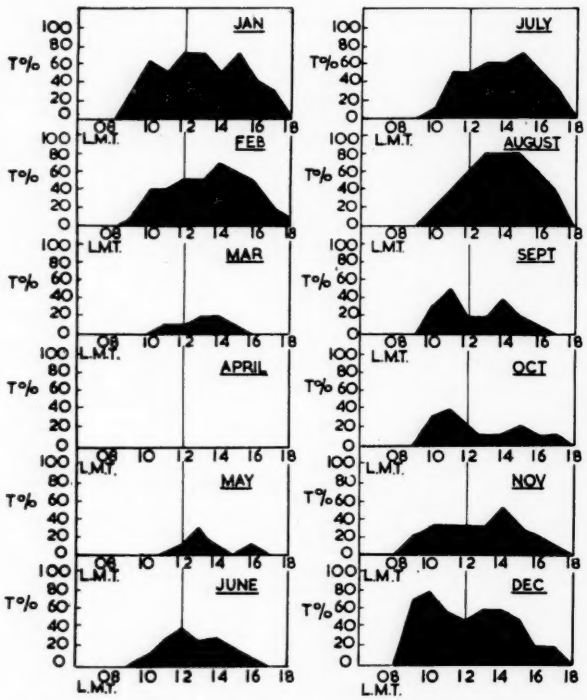


Fig. 2. Diurnal variation of thick layer occurrence at Singapore for each month of the year 1949.

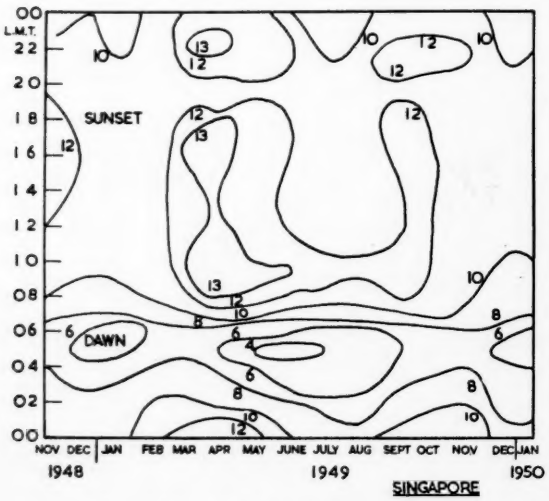


Fig. 3. Contour map of F_2-Layer critical frequency.

The lowest value of cos χ, where χ is the solar zenith angle, for Singapore at noon during any season of the year is about 0·93, so that "summer" conditions are maintained continuously.

The large diurnal variations obtained are, at some periods of the year, similar to those observed at Huancayo, and are unlike the summer behaviour of the layer at higher latitudes. It will also be noticed that the seasonal variation follows the sun, instead of the inverse effect which is found to hold at high latitudes.

It is very significant that, during the equinoctial periods when thick layer phenomena are least evident, the diurnal and seasonal variations of f_0F_2 at Singapore during daylight hours are in good agreement with those obtained at Huancayo, whereas when they are present the observations at the two stations do not agree. This is illustrated in Figure 4, where the diurnal variations of f_0F_2 at Huancayo and at Singapore are shown for the months of January, April and July 1949. The absence of the characteristic 1000 L.M.T. equatorial maximum on the Singapore curve for January, during the presence of the morning thick layer effects, is remarkable. The behaviour at the two stations is very similar between sunrise and sunset during the equinoctial month of April. In July, normal equatorial variations of critical frequency are found at the two stations during the hours before noon, but the afternoon thick layer phenomena give relatively low values of f_0F_2 at Singapore.

It will also be observed that whereas in January Huancayo critical frequencies are higher than those at Singapore, the reverse is true in July. This appears to be due to the presence of a dominant six-monthly period in the seasonal variation of F_2 layer critical frequency at Singapore at noon, compared with an annual cycle at Huancayo where f_0F_2 reaches a minimum in July.

The night-time variation of the critical frequency is very similar to that previously reported for Huancayo, though the exact values found often differ appreciably.

It is concluded that the thick layer phenomena vary with position in the equatorial zone.

4. Scatter Occurrence and Virtual Height Variations in the F Region after Sunset

Large and rapid changes in the minimum virtual height of the F_2 layer after sunset [2], [3] have been found at Huancayo, the virtual height reaching a maximum shortly before 2000 hrs local time. This height increase was generally followed by the appearance of diffuse scatter, observable over a wide range of virtual heights, and often persisting for several hours during the night. In the absence of any well-defined layer, considerable irregularities in the apparent maximum electron density were recorded from hour to hour. This behaviour has been discussed by Booker and Wells [9], who considered the effects to be due to spatial irregularities in the electron density distribution at or above a definite level in the F region, giving Rayleigh scattering. The maximum electron density of the electron clouds necessary to produce the effects shown on the Huancayo records was found to be not more than two or three times the average F layer density.

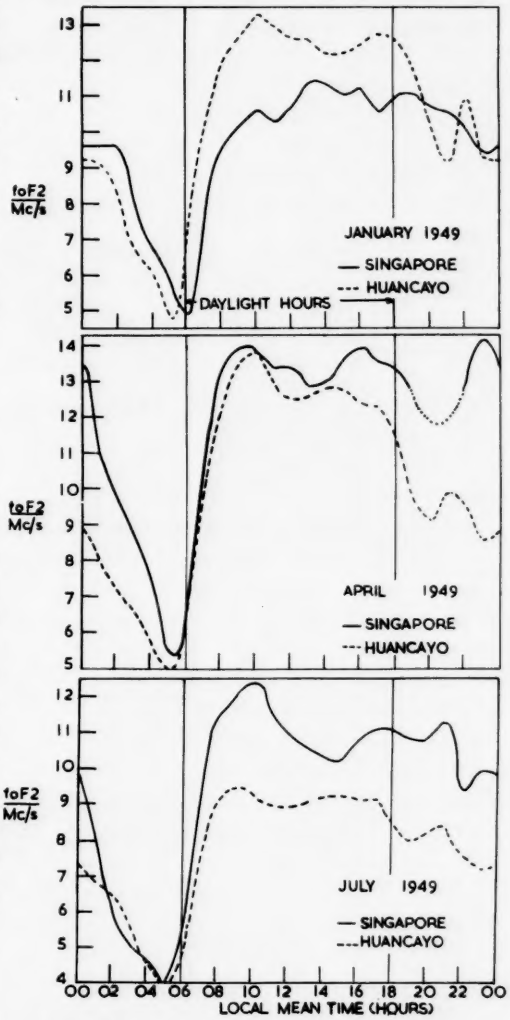


Fig. 4. Diurnal variation of f_0F_2 at Singapore and Huancayo.

The F layer behaviour at Singapore after sunset was often similar. Throughout the year, the minimum virtual height of the layer increased regularly after sunset, reaching a maximum height between 1900 and 2000 L.M.T. The most rapid height

VOL. 2 1952

changes were found at the equinoxes, and the minimum increase in height during June and July.

After reaching its maximum height, the layer often disintegrated into cloud formation, giving a diffuse scattered echo-pattern over wide ranges of virtual height and transmitted frequency. The scatter was observed, at times, up to a maximum virtual height of about 700 km. This corresponds approximately to the maximum height of ionization observed in the hours of daylight when thick layer phenomena were present. When scatter was present, echoes at all heights showed fast fading, the period of the fastest fading cycle being less than one second. It may be mentioned that, though scatter can be recognized on a normal photographic record, the characteristics of the echo-pattern are far more vividly portrayed when observed visually.

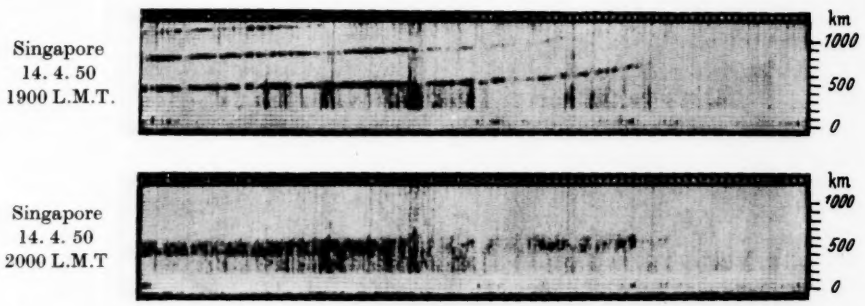


Fig. 5. *h'f* records showing formation of equatorial scatter after sunset.

A diffuse echo-pattern, returned from a wide range of virtual heights, has frequently been obtained at frequencies from 2 Mc/sec up to as high as 13 Mc/sec, without any indication of layer formation, and with rapid and violent fading taking place oyer the entire virtual height range.

A typical example of this layer disintegration is shown by the two records of Figure 5, taken at successive hours during the formation of scatter. The minimum virtual height of the layer which was about 300 km at 1800 L.M.T., has reached 365 km at the time of the first record (1900 L.M.T.), and the layer has completely disintegrated an hour later.

The nature and extent of the scattered echoes varied considerably. On several occasions, echoes from several distinct layer heights were simultaneously obtained at low frequencies, although a clear single reflection was present at higher frequencies. On one such occasion (3 March 1949) information* of reception reports from Siam of a sender situated at Singapore indicated that clear reception was obtained only at relatively high frequencies. This seemed to indicate that at lower frequencies, signals suffered successive reflections between the various cloud formations in the *F* region, with consequently many possible path lengths for a particular range. Similar conditions are very seldom experienced at high latitudes.

At other times, diffuse scatter was only observed above a particular frequency, the cloud formation being formed near, or slightly below, the maximum of the layer. Clear echoes were obtained at low frequencies from the undisturbed lower part of

* Made available by the British Far Eastern Broadcasting Service.

the layer. Such a transition frequency was generally observed to be between 7 Mc/sec and 9 Mc/sec, the most frequently observed range of critical frequencies of the ledges of the thick daytime F region. This is appreciably lower than the range of evening critical frequencies obtained on the days when scatter was absent. This high level cloud effect has been often observed at Huancayo [3].

The monthly mean variation between the hours of 1700 L.M.T. and 2200 L.M.T. of the minimum virtual height of the F_2 layer, $h'F_2$, during October 1949 (a month of typical equinoctial conditions) is shown in Figure 6. It will be observed that $h'F_2$ increases rapidly, after 1800 L.M.T., to a sharp maximum, which is almost a discontinuity. After this maximum, the minimum virtual height of the layer (or of the lowest part of the scattered cloud formation) decreases at a progressively slower rate. As the layer re-forms into a simple structure, the minimum virtual height falls from about 390 km at 1915 L.M.T. to 220 km at 2200 L.M.T. The variation of scatter occurrence, expressed as a percentage of total useful observations, is also shown in Figure 6 on the same time scale.

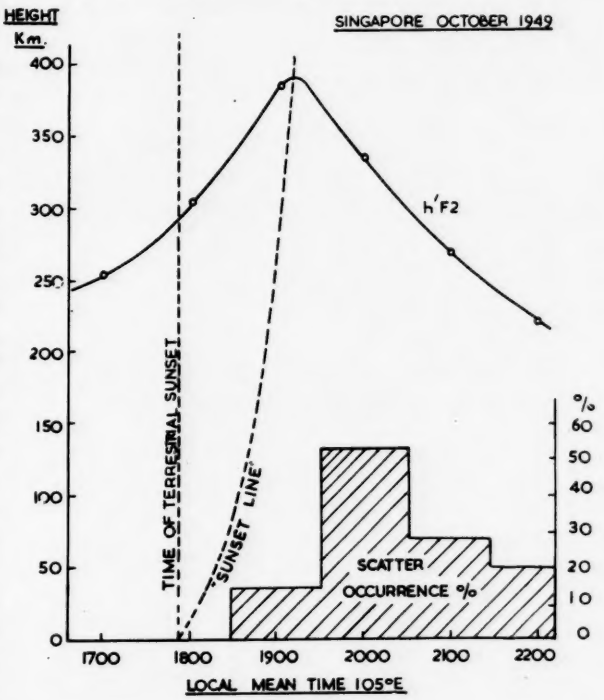


Fig. 6.

The time of sunset on the earth's surface at Singapore in mid-October 1949 was approximately 1752 L.M.T. The variation of the time of sunset with height above the earth's surface is shown by the "sunset line" in Figure 6. This curve was obtained on the assumption that the sun set vertically (which is valid during equinoctial months in equatorial regions), by calculating the time interval after terrestrial sunset from simple geometrical considerations, refraction being neglected.

This "sunset line" was found to intersect the virtual height curve at the virtual height maximum.

This showed that

(a) the greatest increase in layer height occurred between the time of terrestrial sunset and the time of sunset on the layer;

(b) the layer height reached a maximum at the time of sunset on the layer;

and (c) that the disintegration of the layer into scattered clouds generally occurred, as far as can be ascertained from records at hourly intervals, at the time of sunset on the layer.

VOL. 2 1952

The experimental results for other equinoctial months showed similar phenomena, though the number of measurements between 1800 L.M.T. and 2100 L.M.T. was often restricted by power cuts and voltage variations in the local electricity supply.

The mean monthly rate of increase of $h'F_2$ from 1800 L.M.T. to 1900 L.M.T., the period between the approximate times of terrestrial and ionospheric sunset, is shown in Figure 7a. A predominant six-monthly variation in this rate can be observed. The change in the rate of movement of the sun at this time over the year is comparatively small. It will be noticed that the virtual height always increased at this time of day, and that in March 1949 the monthly mean rate of increase exceeded 100 km/hr, an extraordinarily rapid rise for a normal ionospheric layer. The occurrence of scatter at 2000 L.M.T. over the same eighteen-month period, expressed as a percentage of total useful observations, is shown in Figure 7b, and it can be seen that scatter occurrence after ionospheric sunset is associated with rapid height changes during the previous hour.

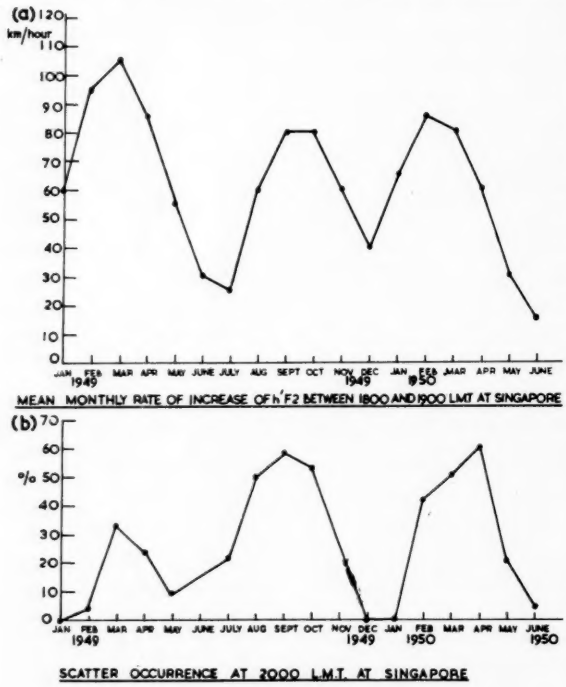

(a)
km/hour
MEAN MONTHLY RATE OF INCREASE OF h'F2 BETWEEN 1800 AND 1900 LMT AT SINGAPORE
(b)
%
SCATTER OCCURRENCE AT 2000 L.M.T. AT SINGAPORE

Fig. 7.

The occurrence of scatter phenomena during the remainder of the night is normally unimportant, apart from the relatively few occasions when the critical frequency is abnormally low.

Scatter occurrence precludes the possibility of determining the maximum usable frequency for a given range: the great increase in layer height, and the subsequent diffuse cloud-like formation of the F region, both tend to cause deterioration in the propagation conditions, and it seems unlikely that normal ionospheric reflection can be relied on for transmission through equatorial regions during the period 1900–2000 L.M.T. at the equinoxes.

5. The Interpretation of the $h'f$ Records at Singapore

5.1 *Definitions and symbols*

The complex F_2 structure which is frequently observed at Singapore cannot be described adequately using the conventional and internationally recognized symbols and terminology.

The abnormal behaviour of the layer leaves some doubt as to whether the highest critical frequency observed should be ascribed to the F_2 layer. The great height at which this ionization often appears also decreases its control of oblique incidence

propagation, and it is possible that, in practice, one of the lower ledges in the layer may control the maximum usable frequency. The height *hm* and the critical frequency of the highest layer show continuity in diurnal variation (though more than one phenomenon may be present) whereas the corresponding parameters for the ledges are usually irregular.

The critical frequency of the layer decreases with increase in the height of the maximum ionization density in a similar manner to that found at higher latitudes, and the highest critical frequency has therefore been taken to be f_0F_2, though the factors controlling the formation of this layer may be different from those of higher latitudes. The critical frequency of ledges comparable with the F_1 layer at higher latitudes could seldom be measured. When present, these ledges were usually very variable in height and density, and did not behave in the regular manner typical of the F_1 layer at higher latitudes. It is concluded that a normal F_1 layer is not present at Singapore. Nevertheless the critical frequencies of these ledges have been quoted as f_0F_1.

The recognized definition of the minimum virtual height of the F_2 layer, "$h'F_2$," is not applicable under thick layer conditions, or when the separate existence of the F_1 layer is uncertain, and the data are therefore not comparable with those obtained in higher latitudes. When the F_2 region is clearly stratified, the highest layer can conveniently be referred to as F_2, and its minimum virtual height is then controlled by the highest ledge. Since these ledges are highly variable in height and density, the values of $h'F_2$ obtained have little significance.

VOL. 2 1952

The most regularly observed minimum virtual height in the F region at Singapore, which was for convenience denoted by "$h'F$" was that for the entire F layer above the F_1 ledge or for the entire F_1 region when this ledge could not be identified. In addition $h'F_1$ was measured whenever a ledge appeared in the normal F_1 region. Further minimum virtual heights for ledges in the higher F layer, which were generally intermittent and irregular from hour to hour and absent for many months, have been measured when possible, and denoted by the symbol "$h'Fh$" (the suffix "h" being the recognized symbol for stratification). The highest value of "$h'Fh$" at any given time would thus correspond to the international definition of $h'F_2$, but it is inadvisable to use this latter symbol under conditions of stratification owing to the possibility of confusion and misinterpretation. For similar reasons all critical frequencies intermediate between f_0F_1 and f_0F_2 have been referred to by the symbol "f_0Fh."

5.2 *The determination of the constants of upper layers*

On the assumption of the existence of a single parabolic layer, the information given by the experimental $h'f$ curve can be used, by the method of Appleton and Beynon [8], [9] referred to above, to obtain values of the height and thickness of the layer. If a lower layer is also in existence, knowledge of its characteristics is necessary in order that any retardation taking place in it can be determined. The $h'f$ curves obtained at Singapore during daytime were seldom such as would be obtained from a compact parabolic layer. On many occasions there were large virtual height gradients at frequencies which were not necessarily close to the highest critical frequency, and there were indications of the possible presence of subsidiary maxima

of electron density, the corresponding critical frequencies being intermediate between f_0F_1 and f_0F_2.

It is well known that virtual height data are not sufficient to determine the electron distribution at heights beyond the first layer maximum. MANNING [10] has given a solution to the virtual height problem, consistent with the experimental curves and obtained by direct integral analysis, using a single-valued distribution of electron density with a point of inflection in the height versus density curve at the electron density corresponding to the critical frequency of the lower layer. This distribution shows no valley between separate layers. If for physical reasons it is believed that several separate layers do exist, then it is possible [11] to find a correction which makes the computed true height include the postulated valley and yet be consistent with the virtual height data. MANNING showed that the width of the assumed valley must be restricted so that the corrected curve of electron density with true height is single-valued, and that, by assuming the largest valley width which still permits interpretation with a physically realizable distribution, a limit is found to the possible error in the determination of the upper layer true height.

It would thus appear to be impossible, using vertical incidence virtual height measurements alone, to obtain a true and unambiguous picture of the ionization distribution of the upper *F* region, at times when complex virtual height curves are observed at Singapore.

Using MANNINGs method [11], applied to the case of a typical equatorial curve obtained under stratification conditions, where the upper critical frequency foF_2 is 11 Mc/sec, the intermediate critical frequency "f_0Fh" (which may or may not be observed as a distinct cusp) is 9 Mc/sec, and the width of the assumed valley is not more than 300 km, the height correction for the upper part of the layer is not more than 200 km for an empty valley, and not more than about 90 km for the much more probable parabolic valley. This type of experimental curve would require the greatest height corrections. For this type of distribution a typical value of *hm* for the upper layer, without correction, would be of the order of 800 km. With allowance for retardation in a parabolic valley, the true height of maximum ionization would be over 700 km. For less extreme distributions giving lower measured values of *hm*, the necessary correction would be less than 90 km.

It would thus seem unlikely that the correction to the apparent height of maximum ionization of the upper *F* layer at Singapore due to the retardation in the lower part of the *F* layer could exceed about 15%. It would therefore seem that, subject to this limited correction, the large fluctuations in the experimental virtual height curves obtained at Singapore have a real meaning, even though the electron distribution at intermediate heights may not be known, and that the true values of the height of maximum ionization density during the morning hours in the December solstice often exceed 700 km.

In conclusion, it can be said that the thick layer ionospheric behaviour at Singapore is very different from that found at high latitudes; is not always similar to that at other equatorial stations; and would be expected to influence local propagation conditions over paths involving reflections within the *F* region.

Some examples of *h'f* records obtained at Singapore in the December solstice during the presence of a thick *F* region are shown in Figure 8. The frequency scale

is given by the white dots above the records, at 20 kc/sec intervals from 2·2 Mc/sec to 7 Mc/sec, and from 7 Mc/sec to 16 Mc/sec, and the horizontal height markers are at 50 km intervals to 1200 km.

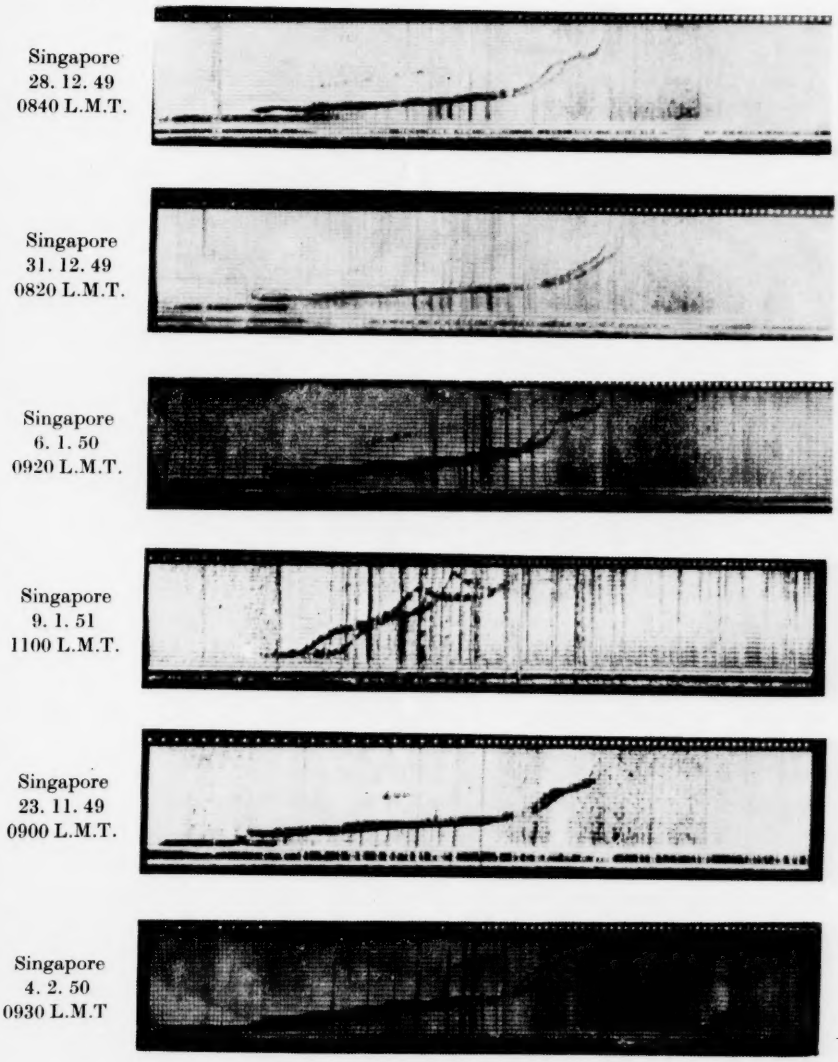


Fig. 8.

Acknowledgements—The work described above was carried out as part of the programme of the *Radio Research Board*. This paper is published by permission of the Director of the *National Physical Laboratory*, and the Director of Radio Research of the *Department of Scientific and Industrial Research*.

Acknowledgements are due to Professor N. S. ALEXANDER, of the *University of Malaya*, for many helpful discussions, and for the facilities made available in the

VOL.
2
1952

Physics Department of the University; and to Mr. A. J. G. MOORAT for his assistance in the analysis of records and in the maintenance of the recorder.

The author is indebted to Mr. W. R. PIGGOTT for guidance in the preparation of this paper.

REFERENCES

[1] U.S. Army; Report on Japanese Research on Radio Wave Propagation, May 1946. [2] WELLS, H.W., and STANTON, H. E.; The Ionosphere at Huancayo, Peru, November and December 1937. Terr. Mag. and Atmos. Elec., 1938 **43** 169. [3] WELLS, H. W. and STANTON, H. E.; Ionospheric Characteristics at Huancayo, Peru, December, 1937, through December, 1938. Terr. Mag. and Atmos. Elec. 1939 **44** 326. [4] HON YUNG SEN; Stratification of the F_2-Layer of the Ionosphere over Singapore. Jour. Geophys. Res. 1949 **54** 363. [5] THOMAS, H. A. and CHALMERS, R. G.; An Improved Ionospheric Height Recorder. J. Inst. Elect. Engrs., Part III, 1948 **95** 7. [6] GILLILAND, T. R.; Ionospheric Investigations. Nature, Lond. 1934 **134** 379. [7] APPLETON, E. V. and BEYNON, W. J. G.; The Application of Ionospheric Data to Radio Communication Problems, Part I. Proc. Phys. Soc. 1940 **52** 518. [8] APPLETON, E. V. and BEYNON, W. J. G.; The Application of Ionospheric Data to Radio Communication Problems, Part II. Proc. Phys. Soc. 1947 **59** 58. [9] BOOKER, H. G. and WELLS, H. W.; Scattering of Radio Waves by the F Region of the Ionosphere. Terr. Mag. and Atmos. Elec. 1938 **43** 249. [10] MANNING, L. A.; The Determination of Ionospheric Electron Distribution. Proc. Inst. Radio Engrs. 1947 **35** 1203. [11] MANNING, L. A.; The Reliability of Ionospheric Height Determination. Proc. Inst. Radio Engrs. 1949 **37** 599.

Journal of Atmospheric and Terrestrial Physics, 1951, Vol. 2, pp. 79 to 82. Pergamon Press Ltd., London

Weitere Ergebnisse der Fade-out-Statistik

DIETRICH STRANZ

(Received 13 June 1951)

ABSTRACT

The investigation of the weak absorption effect of the corpuscular cone emitted by the Sun, detected earlier by the statistical method of sample days, has been continued for the year 1949 yielding a result very similar to that of 1948.

VOL. 2 1952

In einer früheren Mitteilung [6] wurde ein kurzer Überblick über statistische Ergebnisse gebracht, die bei einer Untersuchung von Totalschwunderscheinungen auf Kurzwelle gefunden worden waren. Die Auszählung von Sonnenflecken hat ein scheinbares Defizit um den Zentralmeridian und westlich davon an Tagen mit Kurzwellentotalschwund (SID = sudden ionospheric disturbance) gegenüber Tagen ohne SID gezeigt. Es war versucht worden, die Erscheinung damit zu erklären, daß der Korpuskelstrom [2], [3] von der Sonne zur Erde eine schwache Absorption [1], [2] auf die exzessive UV-Strahlung ausübt, welche SID verursacht. Oder mit anderen Worten: SID machen sich ein wenig bevorzugt an Tagen mit geringerer Fleckenzahl in den kritischen Gebieten der Sonne geltend, d. h. wenn Absorption auf dem Wege von der Sonne zur Erde weniger wahrscheinlich ist.

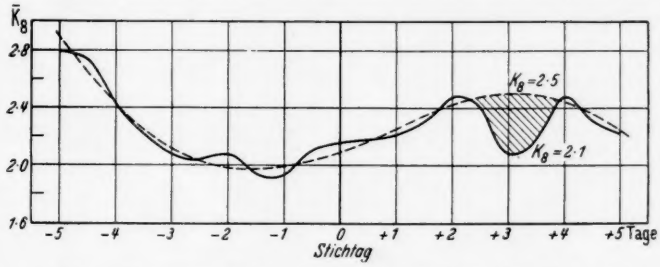


Abb. 1. Durchschnittlicher magnetischer Charakter K (8fach geglättet). 35 sichere SID 1949.

Wie schon anderswo [7] erwähnt, läßt sich eine ähnliche Untersuchung mit Hilfe der magnetischen Kennziffern K (three-hourly range index) anstellen. Davon soll hier berichtet werden. Die früher beschriebenen Versuche (1948) haben ein Beobachtungsmaterial zur Grundlage, das nicht ausschließlich sichere SID-Fälle aufweist. Es ist dennoch schon eine schwache Herausschälung des erwähnten Effektes zu erkennen, der dann mit völlig sicherem Beobachtungsmaterial des gesamten Jahres 1949 wiederholt klarer hervortritt.

Nach der Methode der Stichtage sind die Kennziffern K* des Magnetischen Observatoriums Lovö (Stockholm) für 35 sichere SID-Tage des Jahres 1949, und zwar für die Tage -5 bis $+5$ herausgeschrieben worden. Mit „sicher" ist gemeint, daß sowohl Totalschwund wie solar flare beobachtet sind. Abb. 1 zeigt den durchschnittlichen Charakter des Magnetfeldes für diese 35 Tage. Die K-Werte sind 8fach geglättet, damit der tägliche Gang unterdrückt wird. Wir sehen, daß es 1 bis

* In Zukunft sollten die neuen weltweiten K_p Verwendung finden, um von den Eigenarten einer Station frei zu kommen.

2 Tage vor SID im allgemeinen am ruhigsten ist. Die Aktivität nimmt dann leicht zu, im „Normalfall" etwa nach der gestrichelt eingezeichneten Linie. Und man erkennt am 3. und 4. Tag nach SID ein deutliches Minimum, also zur selben Zeit, zu der die von KIEPENHEUER [4], [5] beschriebene langsame Korpuskelstrahlung die Erde erreichen sollte. Eine sehr schwach angedeutete Einsenkung kann man am 1. Tage nach SID finden. Sie hebt sich jedoch bei nur 35 Fällen nicht genügend aus der Streuung heraus.

Die Einsattelung am 3. und 4. Tage beträgt 0,4 K-Einheiten oder etwa die halbe Amplitude des durchschnittlichen täglichen Ganges.

Man kann die Untersuchung verfeinern. $K=0$ bedeutet, daß die Erde überhaupt nicht von Korpuskeln getroffen wird. $K=0$ ist daher ein besonders geeigneter Indikator, um den schwachen Absorptionseffekt herauszuschälen. SID-Tage sollten nach einem und besonders nach 3 bis 4 Tagen bevorzugt $K=0$ zeigen (Abb. 2).

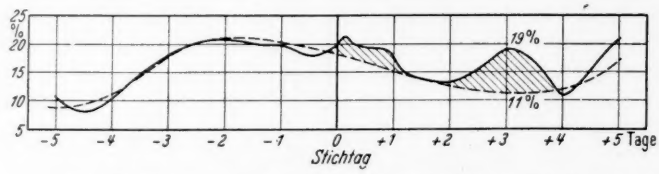


Abb. 2. $K=0$ in % der möglichen Fälle (8fach geglättet).
35 sichere SID 1949.

Der „normale" Verlauf ist antiparallel zu dem der Abb. 1, nämlich viele $K=0$ 1—2 Tage vor SID, abnehmend bis etwa 3 Tage danach, wo aber das ausgesprochene Maximum übernormal ruhigen magnetischen Charakter anzeigt. Das andere Maximum 1 Tag nach SID ist schwächer erkenntlich und wohl nicht so eindeutig (siehe unten).

Am 1. Tag nach SID findet man $K=0$ in etwa 20 anstatt 16% der Fälle, am 3. und 4. Tag in 19 anstatt 11%. Diese Zahlen liegen zwar noch nicht außerhalb des Zufallsbereiches bei der kleinen Anzahl von $n=35$ Fällen im Jahre 1949. Jedoch eine Gegenüberstellung der Ergebnisse der Jahre 1949 mit 35 sicheren Fällen und 1948 (März bis Dezember) mit 130, aber nicht durchweg sicheren Fällen deutet auf einen reellen Hintergrund. 1948 lag noch näher am vergangenen Sonnenfleckenmaximum und weist stärkere Aktivität und damit weniger Fälle $K=0$ auf. Es wurden aus diesem Grunde zum Vergleich beider Jahre die Fälle $K=0$ und $K=1$ zusammengefaßt (Abb. 3a). Der „Normalverlauf" gibt wieder das gleiche Bild mit einem Maximum 1 Tag vor SID und einem Minimum etwa 3 Tage danach. In beiden Jahren überlagert sich ein ausgesprochenes Maximum mit $K=0+1$ drei bis vier Tage nach SID. Im Gegenversuch (Abb. 3b) mit 35 Tagen *ohne* beobachteten Totalschwund und solar flare (1949) ist kein derartiges Maximum vorhanden.

Faßt man schließlich die beiden Jahre 1948 und 1949 zusammen, so daß man $n=165$ Tage zur Verfügung hat, und verwendet die Zeilen $K=0$ und $K=1$ wie vorher, so erhält man Abb. 4, in der das Maximum am 3. und 4. Tag schon so deutlich hervortritt, daß es beinahe mit Sicherheit als überzufällig ($\varepsilon/2=0{,}135\%$) bezeichnet werden kann. Die Zusammenfassung zweier Jahre hat den an sich

schwachen Effekt stärker heraustreten lassen und nicht verwischt, was bei Zufälligkeit eher zu erwarten gewesen wäre.

Das sehr schwache Maximum für $K=0$ am 1. Tag nach SID ist mit zunehmender Zahl der Fälle verschwunden (hier $K=0+1$). Nach Messungen von GARTLEIN [3]

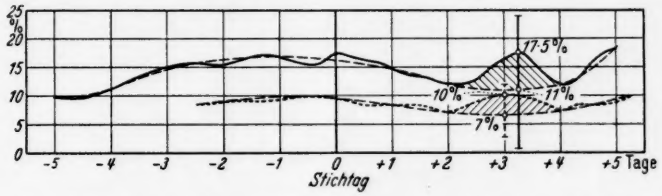


Abb. 3a. $K=0$ und 1 in % der möglichen Fälle (8fach geglättet).
—— 35 sichere SID 1949. ---- 130 SID 1948.

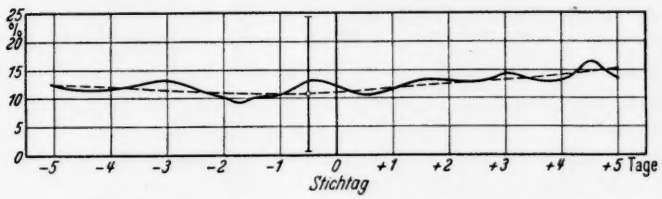


Abb. 3b. $K=0$ und 1 in % der möglichen Fälle (8fach geglättet).
35 analoge Tage *ohne* SID 1949.

und MEINEL (zitiert in [3]) scheint eine Beziehung zu bestehen zwischen Protonenenergie und Nordlichtintensität. Es wurden durchschnittliche Geschwindigkeiten der Protonen von 675 km/sec festgestellt, die maximal bei mäßigem Nordlicht auf 1350 km/sec ansteigen. Bei sehr starkem Nordlicht soll die Geschwindigkeit noch erheblich größer sein.

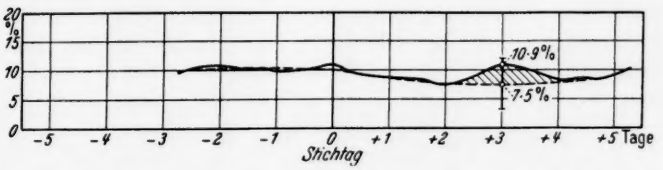


Abb. 4. $K=0$ und 1 in % der möglichen Fälle (8fach geglättet).
165 SID 1948 und 1949.

Die für die Untersuchung ausgewählten SID weisen am 1. Tage nach der Erscheinung 16mal $K=5$ (5,7%), 5mal $K=6$ (1,8%), 1mal $K=7$ (0,4%), 1mal $K=8$ (0,4%) und kein einziges Mal $K=9$ (0%) auf, also insgesamt 8,2% der Fälle haben $K\geq 5$. Dem stehen gegenüber 55mal $K=0$ (19,6%) und 36mal $K=1$ (12,8%) oder 32,4% der Fälle $K=0$ und 1. Auf einen gestörten Tag kommen bei dieser Auswahl somit 4 magnetisch sehr ruhige Tage. Über die Hälfte der Fälle, nämlich rund 60%, sind Tage mit mäßiger Unruhe ($K=2$—4). Nach der von GARTLEIN [3] gefundenen Beziehung sollte also hier vornehmlich die langsame Korpuskelstrahlung zur Geltung kommen, ganz in Übereinstimmung mit dem hier gefundenen Ergebnis, das den Schwerpunkt der Abweichung vom „Normalverlauf" bei etwa 3—4 Tagen

VOL.
2
1952

aufweist, entsprechend einer Geschwindigkeit von etwa 570 km/sec. Die Untersuchungsmethode ist nicht empfindlich für die seltenen kräftigen Protonenemissionen auf der Sonne.

Mit dem Vorrücken in das Fleckenminimum wird die Verfolgung des Phänomens wegen der Abnahme der SID-Fälle schwieriger werden. $K=0$ und 1 wird häufiger vorkommen und das auffällige Maximum in Abb. 4 zurücktreten lassen. Es ist schon ein Unterschied zwischen 1948 und 1949 zu erkennen, trotzdem 1949 verläßlicheres Material zur Verfügung hat.

Dem Direktor der Erdmagnetischen Abteilung im Kgl. Sjökarteverket in Stockholm, Herrn Dr. N. AMBOLT, und Herrn Dr. R. PENNDORF, Needham-Heights, danke ich bestens für die Bereitstellung und Beschaffung von Beobachtungsmaterial.

LITERATUR

[1] BRÜCK, H. A. und RUTLLANT, F.; Some Observations of the *H* and *K* lines in the Solarsspectrum during a Magnetic storm. Mon. Not. Astr. Soc. 1946 **106** 130. [2] CHAPMAN, S.; Solar streams of corpuscles: their geometry, absorption of light and penetration. Mon. Not. Astr. Soc. 1929 **89** 450. [3] GARTLEIN, C. W.; Protons and the Aurora. Nature, Lond. 1951 **167** 277. [4] KIEPENHEUER, K. O.; Eine langsame Partikelstrahlung der Sonne. Naturwiss. 1946 **33** 118. [5] KIEPENHEUER, K. O.; A slow Corpuscular radiation of the Sun. Astrophys. J. 1947 **105** 408. [6] STRANZ, D.; Ergebnisse einer Fade-out-Statistik. A. E. Ü. 1950 **4** 217. [7] STRANZ, D.; The Lowest of the Ionized layers of the Upper atmosphere. Tellus 1950 **2** 150.

Book reviews

ALFVÉN, H.; **Cosmical Electrodynamics.** Pp. 240, with 86 text illustrations, (International Series of Monographs in Physics.) Clarendon Press, Oxford, 1950. 25.— net.

During the last five years cosmic physics has developed with great rapidity. Surprising advances have been made in the study of the cosmological aspect of cosmic rays, while the development of radio astronomy has opened new avenues for the exploration of the sun and the galaxy. All this work bears witness to Alfvén's contention that electromagnetic phenomena are of greater importance in cosmic physics than used to be supposed. The first half of his timely and stimulating book is of a fundamental character and the last three chapters deal with specific applications. Even if some of this later material fails to stand the test of time there can be no question as to the lasting value of his earlier chapters.

After a short survey of magnetic and electric fields in cosmic physics, there is a clear treatment of the motion of charged particles in magnetic fields, including an account of Störmer's work. Workers in diverse fields of physics will find the third chapter on electric discharges in gases of very great interest. The effect of transfering cosmic discharge phenomena to the scale of laboratory experiments is extremely revealing. For example, if one scales down the cosmic phenomena to a laboratory apparatus with linear dimensions of 10 cm then the equivalent pressures are high—the "vacuum" of interstellar space becomes a highly ionized gas at a pressure of 100 atm. The equivalent magnetic fields become enormous—a field of 20 gauss in the solar chromosphere transfers to one of 2×10^9 gauss in the laboratory experiment. The 11 year solar cycle reduces to 0·03 sec and the age of the universe to 100 μ sec! The fourth chapter of the book gives an account of the coupling between the mechanical and electromagnetic forces which produces the magneto-hydrodynamic waves proposed by Alfvén in 1942.

In the remainder of the book, which consists of a long and comprehensive chapter on solar physics, and a shorter chapter on magnetic storms and aurorae and cosmic radiation, Alfven enters highly controversial fields. From the assertion that "the initial cause of sunspots is the convection in the solar core, which after magneto hydrodynamic transmission to the surface causes spots" the author derives a theory of sunspot phenomena, including an explanation of the 11 year cycle. The theory of the corona and prominences is then treated on the basis of the magneto-hydrodynamic theory. Unfortunately, there is little discussion of the radio frequency radiation from the sun. The chapter is important as a presentation of one view of solar phenomena, but its whole basis depends on the existence of the general solar magnetic field and until the experimental information on the existence of such a field becomes more precise it must remain difficult to asses the validity of Alfvén's views. In any case the chapter is of great value if only on account of the accumulation of solar data which is contained in it.

The chapter on magnetic storms and aurorae is useful as a presentation of the discharge theory, but it suffers from a restriction of treatment both as regards the general phenomena and in the treatment of the current alternative views. A strong point in favour of the discharge theory was the inability of the CHAPMAN-FERRARO theory to derive the immediate cause of the aurorae. Recently, however, Martyn has developed the corpuscular theory in this respect and hence controversy between the alternative views has become increasingly vigorous. The chapter on cosmic radiation again suffers from its brevity. It contains a short account of the effect of solar and terrestrial magnetic fields on the radiation, and a rather brief discussion of the origin of cosmic radiation.

There is no doubt that Alfvén has achieved his purpose admirably, namely to present the fundamental features of cosmical magnetic and electrical phenomena and to give a unified view of three specific topics in cosmic physics. Nevertheless one is left with some feeling of disappointment at the restricted treatment of these extremely interesting topics. Even so, the author states that an eighth chapter on an electromagnetic theory of the origin of the solar system has been excluded for reasons of space, and similarly the physics of the ionosphere has been excluded altogether. It will be evident, however, hat no physicist or astronomer can afford to neglect this important book. A. C. B. LOVELL.

BARLOW, H. M. and CULLEN, A. L.; **Micro-wave measurements.** Pp. xvi + 399. Constable and Co., Ltd., London 1950. 30 s. net.

The development and application of very short electric waves during the decade 1940-50 was so rapid and so widespread that it was clearly necessary for those engaged in the work to devise methods of measuring the various quantities involved in the use of the new techniques. Many of the details of these methods have been described in original papers scattered throughout scientific and technical literature, but this book forms a very timely and welcome survey of the present position in this field.

VOL. 2 1952

With regard to the title, the term "micro-wave" is by no means generally accepted and well understood. For example, in his book on "Microwave Transmission" Prof. J. C. SLATER designates the term to cover wavelengths between one centimetre and one metre. In the present case Prof. BARLOW and Mr. CULLEN intend "micro-wave" to indicate the range of wavelengths for which waveguides and resonant cavities are most appropriately used; and the scope of the book under review is practically confined to measuring techniques involving such items.

Introductory chapters dealing with the fundamental characteristics of guided waves, wave impedance, attenuation, and the properties of cavity resonators, provide an admirable introduction to Chapter IV on the measurement of wavelength and frequency. Then the principles of standing wave indicators, and matching and transmission systems are discussed, before proceeding to deal with the measurement of power in Chapter VII. Here, a useful survey is given of the various thermal methods of measuring powers applicable to magnitudes above and below about 10 w. The measurement of attenuation is next discussed: and then the remaining four chapters are devoted to a survey of the application of the various measuring techniques to the study of the electrical properties of materials at very high frequencies, and to the determination of the characteristics of receivers, transmitters, and aerial system. Five appendices extend the theoretical work in the book to certain specific cases of interest.

The treatment of the book throughout comprises a clear description of the principles of measurement of the various quantities, supported by an adequate, although not complicated, analytical discussion of the theory, and an indication of the possibilities and application of the methods involved. The work is illustrated by a large number of very clear line diagrams, but no photographs of measuring instruments are included, nor are there any detailed descriptions of the designs of such instruments. On the other hand, at the end of each chapter is given a list of references: these, amounting to 117 in all, of which 86 are to British publications, and 30 to American, form a most useful and conveniently arranged guide to the literature of the subject of waveguide and measuring technique.

The book is very well produced and will prove a most useful guide and work of reference not only to engineers and scientists concerned with communications, radar and radio-astronomy, but also to the physicist conducting general research with the aid of micro-wave technique.

R. L. SMITH-ROSE.

CORRECTION

Our attention has been drawn to some minor errors which occur in the paper by K. C. WESTFOLD on "The Interpretation of the Magneto-Ionic Theory," published in Vol. 1 of this journal, pages 152–186. In equations (2.35), (2.36) and (2.37) the results for Q_α and Q_β are incorrectly given. These should read as follows:

$$Q_\alpha = i\{g - (|g|/g)\sqrt{(1+g^2)}\}; \qquad Q_\beta = i\{g + (|g|/g)\sqrt{(1+g^2)}\}.$$

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 85 to 105. Pergamon Press Ltd., London

Group velocities and group heights from the magneto-ionic theory

D. H. SHINN* and H. A. WHALE**

(Received 7 July 1951)

ABSTRACT

In Part I the group velocity of a radio wave in the ionosphere above South East England is calculated for the ordinary and extraordinary wave for a wide range of frequencies. Results are presented in the form of curves. A sufficient number of values has been calculated for group delays at vertical incidence to be computed. In Part II similar values for various angles of dip are used to find the shapes of the group height *vs.* frequency curves, for the ordinary ray vertically incident on a parabolic layer, at various magnetic latitudes. Estimates of height of maximum ionisation and thickness of layer from observed group height *vs.* frequency curves have until now been based on calculations in which the effect of the magnetic field is neglected. It is shown that such estimates of layer thickness are correct on the magnetic equator, but yield values which are too high by about 17% at magnetic latitude 27° rising to about 53% at magnetic latitude 62°. Estimates of height of maximum ionisation are, however, approximately correct. A modification to the present method of estimating these parameters is proposed. The effect of the field on oblique incidence propagation is discussed, but no exact conclusions are reached.

VOL.
2
1952

INTRODUCTION

It is well known that the group velocity for both the ordinary and extraordinary waves in the ionosphere depends on the direction and magnitude of the earth's magnetic field in a complicated way. Its behaviour has been previously studied by several authors, and some numerical values have been deduced and presented in the form of curves [2], [3], [4], [5], [6]. It has been found that these curves are not sufficiently complete to be used in numerical calculations of the "effective height" or "group height" of a postulated ionospheric layer as a function of frequency. Such calculations are necessary for deducing the characteristics of ionospheric layers from observations of group heights.

It is the purpose of the first part of this paper to present a sufficient number of values of the group velocity to enable group heights at vertical incidence to be calculated for any distribution of ionisation over South East England. The values are given in terms of a "group refractive index," μ', which is defined as the free space velocity of light divided by the group velocity of the wave. Curves are presented showing the variation of μ' with electron density and frequency when the wave normal makes an angle of 23° 16′ with the earth's magnetic field. The tables from which the curves were prepared are available from either of the authors, and may be used by those who require greater accuracy than is obtainable by reading the curves. In the second part of the paper the group height is calculated as a function of frequency for a parabolic layer at various magnetic latitudes. The results of these calculations are used to show the effect of the earth's magnetic field on the apparent thickness and height of this layer.

In a previous paper WHALE and STANLEY [7] have deduced curves which show how $\mu\mu'$, the product of the group and phase refractive indices, varies with the

* Research Laboratories, Marconi's Wireless Telegraph Co. Ltd., Great Baddow, Essex.
** Radio Research Laboratory, Auckland University College, New Zealand.

electron density and the strength and direction of the magnetic field. Although these curves are valuable for indicating the magnitude of the perturbations due to the magnetic field, they do not provide the detailed accuracy required for the present purpose. In this paper, therefore, the values of μ' are calculated directly.

The numerical computations were performed on the EDSAC electronic computer in the University Mathematical Laboratory, Cambridge.

The term "group velocity" is used throughout this paper to signify the vertical component of the group velocity. There is also, in general, a horizontal component; a wave sent up vertically from the ground acquires a horizontal motion when it enters the ionosphere. The horizontal motion is not treated in this paper. For the purpose of calculating how long a group takes to travel through a layer, or to reach a particular height, only the vertical motion need be considered. The ionosphere is assumed to be horizontally stratified, *i.e.*, the electron density is a function of height only. The effect of electronic collisions is neglected; this affects only slightly the validity of the results obtained.

PART I

1. *The magneto-ionic formula for the group velocity*

APPLETON'S [1] expression for the phase refractive index, μ, in the absence of electronic collisions, may be written

$$\mu^2 = 1 - 2x(1-x)/D \tag{1}$$

where $x = f_0^2/f^2$,
$f_0^2 = Ne^2/\pi m$, proportional to the electron density, N,
f = frequency at which the refractive index is being measured,
e = charge of an electron in e.s.u.,
m = mass of an electron in grams,
$D = 2(1-x) - y^2 \sin^2\Theta + S$,
$y = f_H/f$,
f_H = gyromagnetic frequency $= He/(2\pi m c)$,
H = magnetic field in oersted,
c = velocity of light in free space,
Θ = angle between the wave normal and the earth's magnetic field,
$S = \pm\sqrt{y^4 \sin^4\Theta + 4y^2(1-x)^2\cos^2\Theta}$.

The positive sign refers to the ordinary component and the negative sign to the extraordinary component. We have put the LORENTZ polarisation term equal to zero, as is common practice nowadays.

The group refractive index, μ', as defined in the Introduction, is given by

$$\mu' = \mu + f\frac{d\mu}{df}.$$

Equation (1) may be written

$$(\mu^2 - 1)\, Df^2 = 2x(x-1)f^2. \tag{2}$$

Differentiation with respect to f leads to the equation

$$(\mu\mu' - \mu^2)\, D = 2\left[-\mu^2 - x^2 + 1 + \frac{(1-x^2)(1-\mu^2)\, y^2\cos^2\Theta}{S}\right]. \tag{3}$$

The values of μ' now to be presented were computed from equations (1) and (3) with Θ put equal to 23° 16′ (the complement of the magnetic dip in South East England). In order to exhibit the infinities of μ' it is plotted on a reciprocal scale.

2. *The ordinary component*

Figures 1 and 2 show curves of μ' for the ordinary component plotted against x (proportional to the electron density) for $\Theta = 23° 16'$ and different values of y

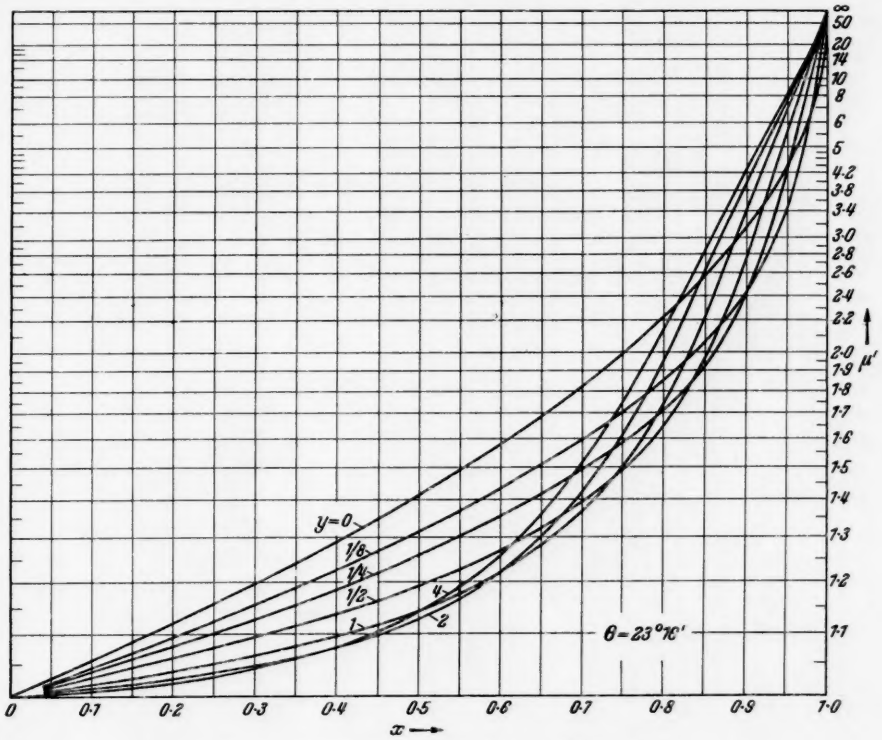


Fig. 1. $\mu' = c/(\text{group velocity})$, for ordinary ray.

$(= f_H/f)$. When $y = 0$ (infinitely great frequency)

$$\mu\mu' = 1 \quad \text{and} \quad \mu' = \frac{1}{\sqrt{(1-x)}}$$

so that, when μ' is plotted against x on a reciprocal scale, the curves take the form of the parabola labelled $y=0$. This is, of course, identical with the variation of μ' with x when there is no magnetic field.

For another frequency, such as that represented by $y = \frac{1}{2}$ $(f = 2f_H)$, μ' is seen to be less than $1/\sqrt{(1-x)}$ for small values of x, but greater than $1/\sqrt{(1-x)}$ for larger values of x. Let us consider a signal travelling up into an ionised medium in which the ionisation density increases until it is sufficient to reflect the wave. It will travel more quickly in the lower part of its path, and more slowly in the upper part of its

VOL. 2 1952

path, than it would in the absence of a magnetic field. The effect of the magnetic field on the group delay will thus depend on the relative times spent in the two regions. This has been pointed out previously by MILLINGTON [4].

When x approaches unity it is convenient to express the formula for μ' as a power series by writing

$$1-x=\varepsilon \text{ (very small)},$$

giving

$$\mu'=\frac{\operatorname{cosec}\Theta}{\sqrt{\varepsilon}}\left[1-\frac{3}{2}\varepsilon\cot^2\Theta+\varepsilon^2\cot^2\Theta\left(1+\left(\frac{15}{8}-\frac{5}{2y^2}\right)\cot^2\Theta\right)\right]$$

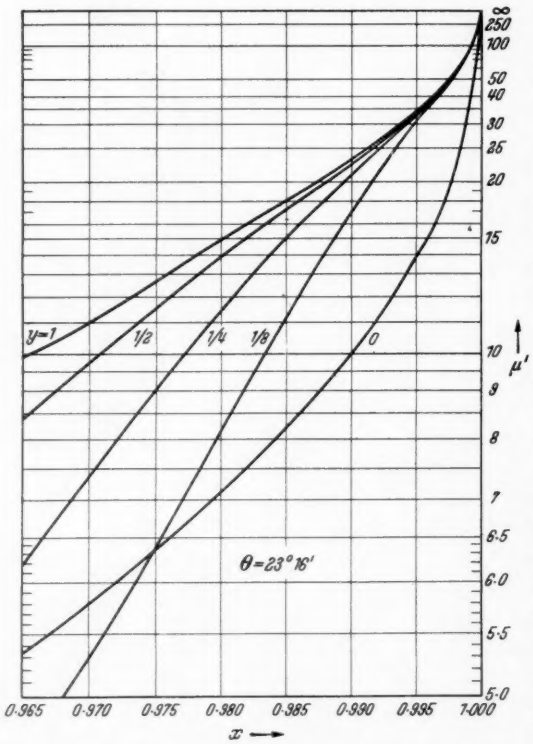


Fig. 2. μ' for ordinary ray, near the reflection point.

Values of μ' when x approaches unity and y is less than unity have been plotted on a larger scale in Figure 2.

3. The extraordinary component, $y<1$

It is well known (*e.g.* RATCLIFFE [17]) that when $y<1$ (*i.e.* for frequencies above the gyrofrequency) the extraordinary wave can be propagated when the electron density has values corresponding to values of x between 0 and $1-y$, or between $(1-y^2)/(1-y^2\cos^2\Theta)$ and $1+y$. The second range corresponds to an unusual type of wave, and will be dealt with in the Appendix; in this Section we will consider only propagation in the range $0<x<1-y$. In this case $x=1-y$ corresponds to the critical density, and it is convenient to plot μ' (on a reciprocal scale) against $x/(1-y)$ as in Figure 3. For $y=0$ (*i.e.*, infinite frequency or zero magnetic field) the curve is the same as that for $y=0$ in Figure 1. All the curves lie entirely above that for $y=0$, *i.e.* the group velocity in the presence of a magnetic field for an electron density $N(1-y)$ is always less than that with no field for an electron density N. It follows that the group velocity in the presence of a magnetic field is always less than that with no field for the same electron density. The dotted curves are for $\Theta=0°$; they show that the group velocity of the extraordinary wave, although it varies considerably with the strength of the magnetic field, is almost independent of its direction for small values of Θ (up to at least 23°). This gives a further justification of the assumption often made (*e.g.* RYDBECK [18], WHALE [19]) that the extraordinary wave behaves at medium and high latitudes in the same way as it does at the magnetic poles, for frequencies greater than the gyrofrequency.

VOL. 2 1952

4. Group delay as a function of frequency for ordinary and extraordinary components

Once the values of μ' as a function of x and y have been calculated, they can be used to calculate the values of group delay, h', as a function of frequency, f, for the ordinary and extraordinary waves propagated through a given ionospheric layer; it is interesting to enquire what differences we might expect between these curves. As a first step towards answering this question, (h', f) curves have been calculated

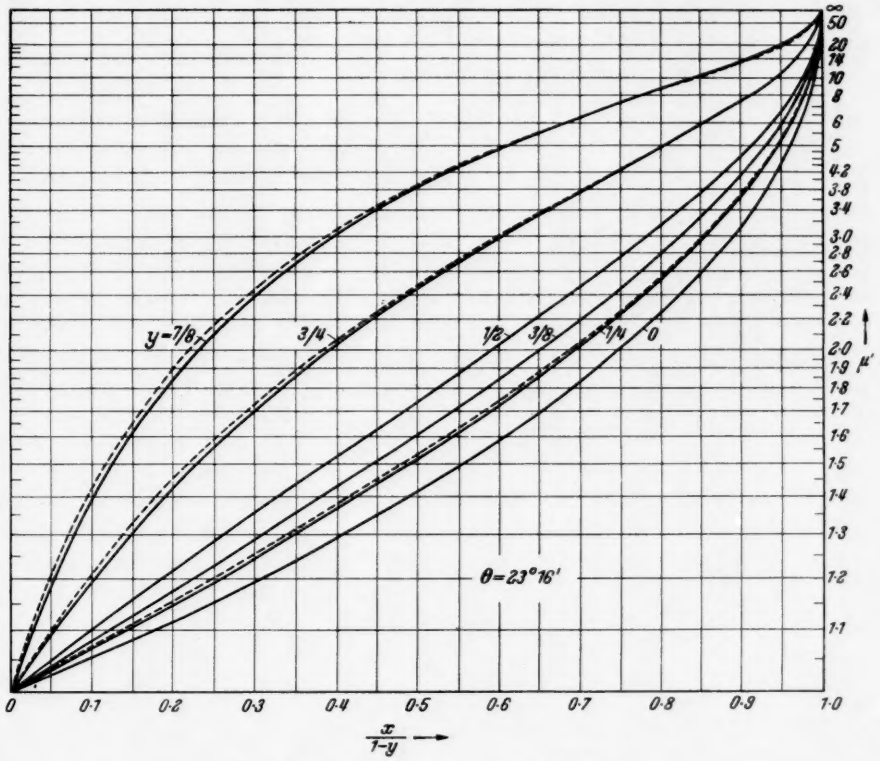


Fig. 3. μ' for extraordinary ray, when $y < 1$.

for a region of uniform electron density. Figure 4a shows the (h', f) curves which would be observed if a uniform layer of unit height with a penetration frequency $f_o = 2$ Mc/sec ($f_H = 1{\cdot}2$ Mc/sec) lay below a sharp totally reflecting layer. They show the total group retardation suffered by the wave in a single traverse of the lower region. In order to compare the shapes of the curves, the (h', f) curve for the extraordinary wave has been drawn both in its correct position, and with a displacement along the frequency axis such that its infinity coincides with that for the ordinary wave. There is also shown, as a dotted line, the (h', f) curve which would be observed if no magnetic field were present. It is interesting to note that the curve for the ordinary wave is distorted at least as much as that for the extraordinary, but in the opposite direction.

It is interesting to compare the calculated shapes of the two curves with the shapes obtained experimentally. Figure 4b shows some experimental (h', f) curves

for the reflection from Region F of waves of frequencies just greater than the penetration frequency of Region E. The extraordinary traces have been shifted to the left by the calculated frequency separation. The shapes of the experimental curves agree well with those of the theoretical curves calculated for the simplified model. It can be shown that the general shape of the curves it not very sensitive to the distribution of ionisation in the lower layer.

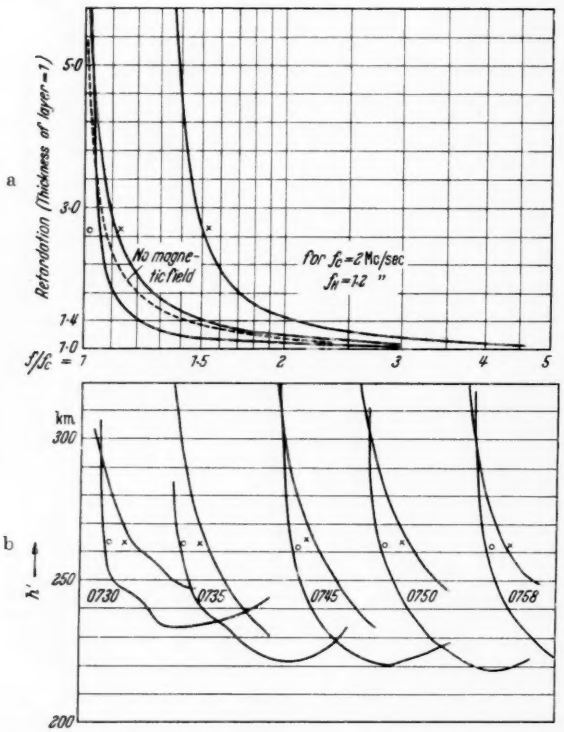


Fig. 4a. Theoretical retardations of the ordinary and extraordinary components on passing through a "rectangular" layer of electrons—the critical frequencies have been adjusted for comparison.

Fig. 4b. Experimental F-region retardations, 16 February 1950, Cambridge.

5. *The extraordinary component, $y>1$*

If y is greater than unity (*i.e.* for frequencies below the gyrofrequency), the critical electron density is reached when $x=1+y$, and it is convenient to show curves of μ' (on a reciprocal scale) plotted against $x/(1+y)$ as in Figure 5. Here, when the reflection level is approached, the signal travels more quickly [for an electron density $N(1+y)$] than in the no-field case (for an electron density N); this is in contrast to its behaviour in the two previously considered cases. When $x=1$ the value of μ' may be easily calculated from the relation

$$\mu'_{x=1} = 1 + \frac{1}{y^2 \sin^2 \Theta}.$$

The locus of such points for $\Theta = 23° 16'$ is plotted as the dotted curve in Figure 5. This gives an easily calculable point near the maximum of the peak on each curve.

PART II

1. *Group height as a function of frequency for a parabolic layer*

The values of the group velocity may be used for computing the (h', f) curves for any chosen ionisation distribution and magnetic field. Some results for a linear layer have been calculated by GOUBAU [3] and POEVERLEIN [5]. Here we present curves for a parabolic layer, which we shall compare with the corresponding curves when the effect of the earth's magnetic field is neglected (APPLETON [8], APPLETON and BEYNON [9], [10], BOOKER and SEATON [11]). We shall calculate the effect of the magnetic field in different parts of the world by using, in turn, the field parameters appropriate to several different magnetic latitudes.

The layer is defined by

$$1-N/N_o=(1-z/z_o)^2, \tag{4}$$

where z = height above lower boundary of layer,
z_o = height of maximum electron density above lower boundary of layer,
N = number of electrons per cubic centimetre,
N_o = maximum electron density.

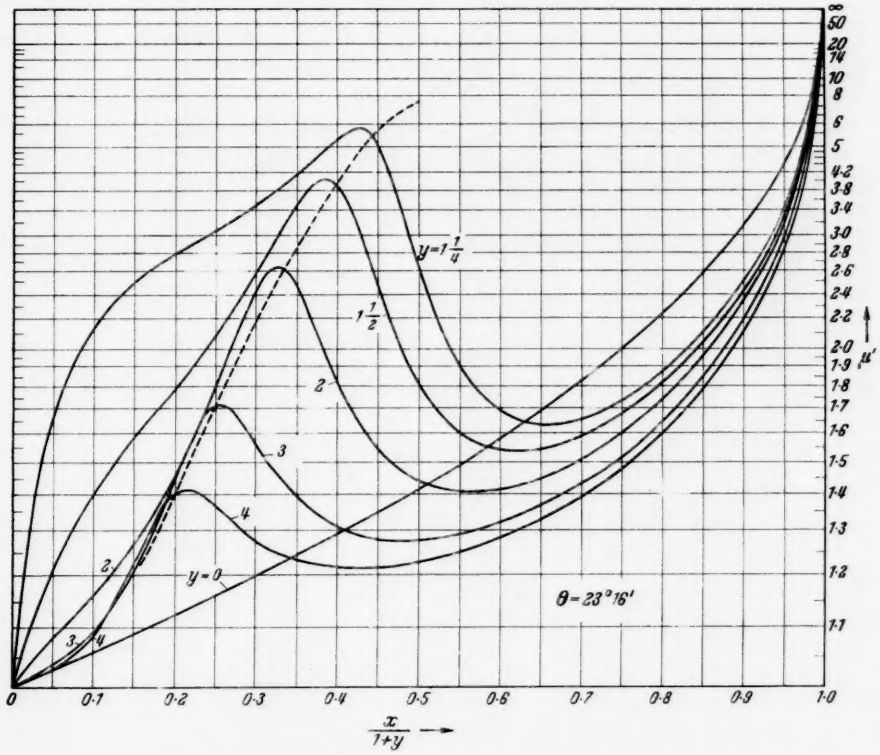


Fig. 5. μ' for extraordinary ray, when $y > 1$.

z and z_o correspond exactly to APPLETON and BEYNONs y and y_m respectively. The critical frequency of the layer is f_c, where $f_c^2 = N_o e^2/\pi m$. From the relations given in Part I, Section 1, we deduce that $N/N_o = x f^2/f_c^2$. The group height, h', is given by

$$h'(f) = \int_0^{z_r} \mu' dz + h_o,$$

where h_o = height of lower boundary of layer above ground level, and z_r = height of point of reflection of frequency f above lower boundary of layer.

From equation (4) we obtain

$$dz = \frac{1}{2} z_o (1-N/N_o)^{-\frac{1}{2}} \frac{dN}{N_o}$$
$$= \frac{1}{2} z_o (1-x f^2/f_c^2)^{-\frac{1}{2}} \frac{f^2 dx}{f_c^2}$$

The transformation $x = 1 - t^2$ now yields

$$h'(f) = h_o + \frac{f^2 z_o}{f_c^2} \int_0^1 \frac{\mu' t \, dt}{\sqrt{(1 - x f^2/f_c^2)}}. \tag{5}$$

The purpose of the transformation is to make the integrand finite throughout the range of integration*.

In the absence of a magnetic field, $\mu' = 1/t$; the integration in equation (5) can be performed easily, and leads to

$$h'(f) = h_o + \frac{f z_o}{2 f_c} \log_e \frac{f_c + f}{f_c - f}. \tag{6}$$

This equation was first given by APPLETON [8], who also showed that the parameter z_o of the parabola which is the best approximation to a simple CHAPMAN layer is equal to twice the scale height of the atmosphere. Equation (6) is often used to deduce the height of maximum ionisation and the thickness of the parabolic layer from observed (h', f) records. It is particularly interesting to investigate how the validity of these deductions is affected by the presence of the earth's magnetic field.

We assume that the field is due to a dipole of moment M at the centre of the earth. We denote magnetic latitude by λ, distance from the centre of the earth by r, the horizontal and vertical components of the field by H_h and H_v, and the angle of dip by φ. Then

$$\begin{aligned} H_v &= 2M \sin \lambda / r^3 \\ H_h &= M \cos \lambda / r^3. \end{aligned}$$

$$\text{Total field} = H \propto \sqrt{(1 + 3 \sin^2 \lambda)} \propto \sqrt{1/(4 - 3 \sin^2 \varphi)} \tag{7}$$

$$\begin{aligned} \sin^2 \varphi &= 4 \sin^2 \lambda / (1 + 3 \sin^2 \lambda) \\ \sin^2 \lambda &= \sin^2 \varphi / (4 - 3 \sin^2 \varphi). \end{aligned}$$

The values of H and φ at the ground at Abinger, Surrey, in January 1950, were 0·47115 oersted and 66° 44′ respectively. The gyromagnetic frequency is given by $f_H = He/2\pi mc$. Table 1 shows the gyromagnetic frequencies at different heights vertically above Abinger; these were calculated from the above formulae assuming that the radius of the earth is 6370 km.

Table 1

Height (km) .	0	120	200	300	400	600
f_H (Mc/sec) . .	1·319	1·247	1·202	1·149	1·099	1·007

Selected values of φ are shown in Table 2, together with the corresponding values of λ, H at ground level, f_H at a height of 300 km, and values of $(h'(f) - h_c)/z_o$ for several frequencies when $f_c = 7$ Mc/sec. Equation (7) was used to calculate the values of H and λ for each selected value of φ, using the values of H and φ at Abinger as a standard. The values of H thus obtained may be in error (since the earth's field is not exactly that of a dipole), but we can allow for this. As it is only the ratio

* It was introduced by MILLINGTON [4], who also introduced other transformations to deal with the extraordinary ray, penetration through the layer, and different shapes of layer.

f_H/f_c which appears in the formulae, the results apply to any other case in which this ratio is the same. For example, suppose that at a point on the earth's surface $\varphi = 45°$ and $H = 0{\cdot}397$ oersted, instead of $0{\cdot}3611$ oersted as computed from equation (7). Then the results of the column headed 45° apply to a critical frequency of 7·7 Mc/sec

Table 2. (For $f_c = 7$ Mc/sec)

φ	0°	45°	60°	66° 44′	75°	81°
$\sin^2\varphi$	0	0·5	0·75	0·8438	0·9330	0·9755
H at ground level (oersted)	0·2855	0·3611	0·4316	0·4712	0·5210	0·5510
f_H at 300 km (Mc/sec)	0·6962	0·8807	1·0525	1·149	1·271	1·344
λ (approx.)	0°	27°	41°	49°	62°	72°
$10^5 . \dfrac{h' - h_o}{z_o}$ for $f/f_c = 0{\cdot}5$	2746·5	2915	3016	3067	3144	
For $f/f_c = 0{\cdot}75$	7297·2	7714	7962	8098	8311	
For $\dfrac{32f}{f_c} = 25$	8192·0	8665	8951	9110	9353	
For $\dfrac{32f}{f_c} = 26$	9216·5	9762	10093	10280	10562	
For $\dfrac{32f}{f_c} = 27$	10412·3	11050	11444	11666	12001	
For $\dfrac{32f}{f_c} = 28$	11847·7	12613	13093	13368	13775	
For $\dfrac{32f}{f_c} = 29$	13649·3	14601	15215	15570	16084	
For $\dfrac{32f}{f_c} = 29{\cdot}5$	14762·7	15845	16558	16974	17571	
For $\dfrac{32f}{f_c} = 30$	16096·8	17354	18208	18713	19430	
For $\dfrac{32f}{f_c} = 30{\cdot}5$	17774·4	19293	20358	20995	21890	
For $\dfrac{32f}{f_c} = 31$	20068·3	22009	23436	24304	25539	
For $\dfrac{32f}{f_c} = 31{\cdot}5$	23842·4	26651	28917	30350	32411	
For $\dfrac{32f}{f_c} = 31{\cdot}75$	27489·8	31340	34721	36955	40279	

SIMPSONS Rule with 17 ordinates is too inaccurate to be used here; see Appendix.

at a height of 300 km, or to a critical frequency of 8·06 Mc/sec at a height of 200 km, and so on. However, we will assume in what follows that equation (7) is exact, and that the values of H shown in Table 2 are correct.

The values of $(h'(f) - h_o)/z_o$ were computed from equation (5) using SIMPSONS Rule with 17 ordinates. The complete computation was carried out on the EDSAC; each value of $(h'(f) - h_o)/z_o$ was computed twice by slightly different methods in order to provide a check; the total machine time required was about three hours. Each individual value of the integrand in equation (5) which was used in the computation was recorded; this information will be useful in investigating the dependence of group velocity on the angle Θ between the wave normal and the field.

VOL. 2 1952

We have assumed that the magnetic field is constant throughout the layer, which is not true. We must now specify the height in the layer at which the field should be measured. In order to estimate the best height to choose, and also the error caused by the assumption, let us consider a practical case in England. We will suppose that $h_o = 225$ km and $z_o = 75$ km. The height of reflection of the frequency $0{\cdot}9\, f_c$ is then 267 km. It may be argued from Figure 1 and Table 2, and it is also evident from the individual values of the integrand in equation (5), that the magnetic field has its greatest effect on the group path when x lies between 0·95 and 1·0. Therefore, if we calculate the group path assuming that H is constant throughout the layer, we

Table 3. (For $f_c = 3{\cdot}5$ Mc/sec)

φ	45°	60°	66° 44′	75°	81°
$10^5 . \frac{h'-h_o}{z_o}$ for $f/f_c = 0{\cdot}5$	2987	3143	3222	3320	3422
For $f/f_c = 0{\cdot}75$	7910	8307	8510	8788	9042
For $\frac{32f}{f_c} = 25$	8889	9344	9578	9902	10187
For $\frac{32f}{f_c} = 26$	10017	10544	10817	11200	11524
For $\frac{32f}{f_c} = 27$	11345	11967	12292	12753	13126
For $\frac{32f}{f_c} = 28$	12956	13709	14108	14680	15123
For $\frac{32f}{f_c} = 29$	15011	15961	16472	17216	17772
For $\frac{32f}{f_c} = 29{\cdot}5$	16300	17393	17988	18863	19505
For $\frac{32f}{f_c} = 30$	17866	19157	19868	20929	21696
For $\frac{32f}{f_c} = 30{\cdot}5$	19872	21453	22346	23690	24660
For $\frac{32f}{f_c} = 31$	22682	24752	25954	27812	29166
For $\frac{32f}{f_c} = 31{\cdot}5$	27470	30596	32522	35634	38001
For $\frac{32f}{f_c} = 31{\cdot}75$	32271	36718	39623	44548	48543

should take the value of H corresponding to $x = 0{\cdot}975$. This corresponds to a height of 265 km for $f = 0{\cdot}9\, f_c$, and 283 km for $f = f_c$. The change in H between these two heights is 0·8 %. Since it takes a change of 50 % in f_H/f_c to cause a change of from 5 % to 10 % in $(h'-h_o)/z_o$ (see Table 3), the error in this last quantity due to the assumption that H is constant throughout the layer is about 0·1 % or 0·2 %. This applies to values of f greater than $0{\cdot}9\, f_c$; H should be taken as the value of the magnetic field at a height $h_o + 0{\cdot}65\, z_o$. Thus the results in Table 2 apply (with an accuracy of 0·1 % or 0·2 %) to a critical frequency of 7 Mc/sec when $h_o + 0{\cdot}65\, z_o = 300$ km.

The complete computation has been repeated for a critical frequency of 3·5 Mc/sec at a height of 300 km; this would apply also to a critical frequency of 3·8 Mc/sec at a height of 120 km, a typical E-layer condition. The results of the second computation are presented in Table 3, and the results of both computations in Figures 6 and 9.

Further information on the gyrofrequency and angle of dip at different points on the earth is presented by BOOKER [20]. His values are incorrect for England ($\Theta = 70°$ instead of 66° 44′,

and f_H at 100 km = 1·42 Mc/sec instead of 1·26 Mc/sec), but they are probably better over the whole earth than those deduced by our method would be. He takes the earth's magnetic field as that due to a dipole of moment $8{\cdot}06 \times 10^{22}$ ampere metre^2 in the direction from 78·6° N, 70·1° W, to 78·6° S, 109·9° E. The correction discussed above ($\varphi = 45°$, $H = 0{\cdot}397$ oersted) corresponds approximately to BOOKER's information. If true, it makes 1% or 2% difference to the values in Tables 2 and 3, and almost no difference to the conclusions of Section 2.

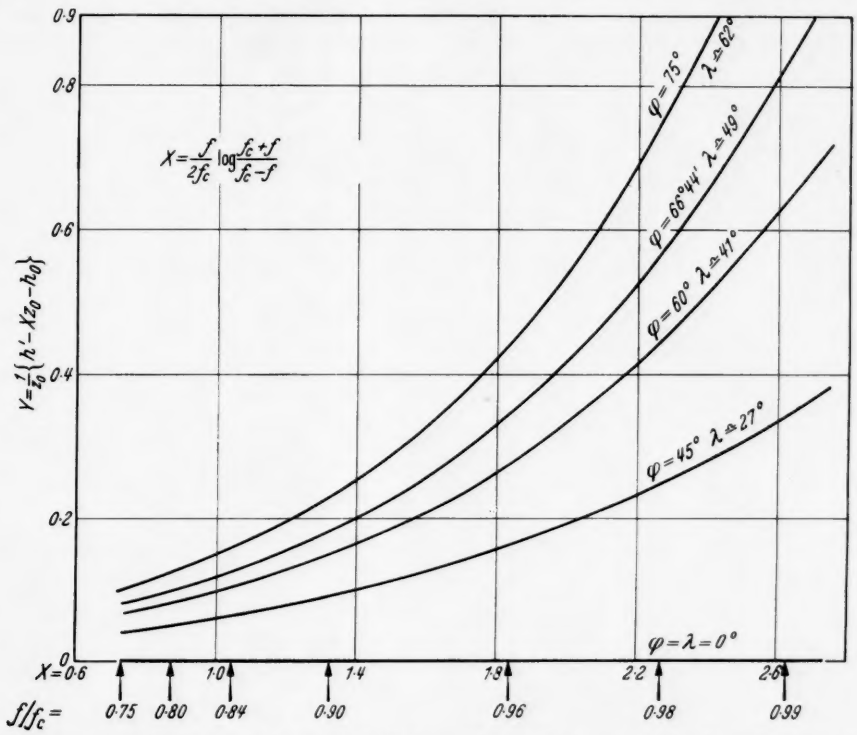


Fig. 6. Computed (h', X) curves for a parabolic layer of semi-thickness z_0. Angle of dip = φ. Geomagnetic latitude = λ. Critical frequency = f_c = 7 Mc/sec. Height of lower boundary of layer = h_0.

2. *Errors in apparent height and thickness of a parabolic layer*

The normal method of deducing h_o and z_o of an ionospheric layer from an observed (h', f) curve is to use equation (6), in which the magnetic field is neglected. It is evident from this equation that estimates of z_o and h_o may be obtained by plotting the observed h' against X, where $X = (f/2f_c) \log \{(f_c+f)/(f_c-f)\}$, and noting the slope and intercept on the h' axis of the resultant straight line (APPLETON and BEYNON (9)]. Due to the effect of the magnetic field this process will yield incorrect values of h_o and z_o. Let us consider the parabolic layers discussed in the previous section, and apply the above process to the computed (h', f) curves. We thus find the "apparent values" of h_o and z_o; these we denote by η_o and ζ_o. The height of the maximum of the layer is denoted by h_m, its "apparent value" by η_m. The departures of the apparent values of h_m and z_o from their real values are of importance, both for the investigation of the characteristics of the atmosphere, and for the prediction of maximum usable frequencies for propagation at oblique incidence.

In Figures 6 and 9 we have plotted the computed values of $(h'-h_o-z_oX)/z_o$ against X. In the absence of a magnetic field these curves would be straight lines along the X axis. Departures from this axis indicate the effect of the field. Our "apparent values" of h_o and z_o are obtained by drawing a straight line through the (h', X) points. We will take as an approximation to this the tangent to the (h', X) curve. The apparent values of h_o and z_o will depend on the point at which we choose to draw the tangent. It may be shown that

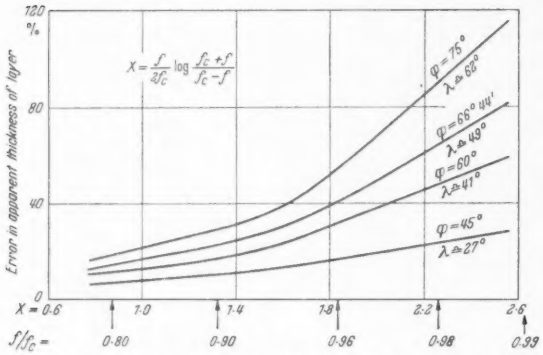


Fig. 7. Error in apparent thickness of a parabolic layer. Critical frequency = 7 Mc/sec.

$$\zeta_o = z_o\,(1 + d\,Y/d\,X)\,, \qquad (8)$$

where

$$Y = \{h' - h_o - z_o X\}/z_o\,.$$

Also

$$\eta_o = h' - X\,\zeta_o\,.$$

Hence

$$\eta_m = h' - (X-1)\,\zeta_o\,.$$

Therefore

$$\left.\begin{aligned}\eta_m - h_m &= h_o + z_o X + z_o Y - (X-1)\,\zeta_o - h_o - z_o \\ &= z_o\,[\,Y - (X-1)\,d\,Y/d\,X\,]\,.\end{aligned}\right\} \qquad (9)$$

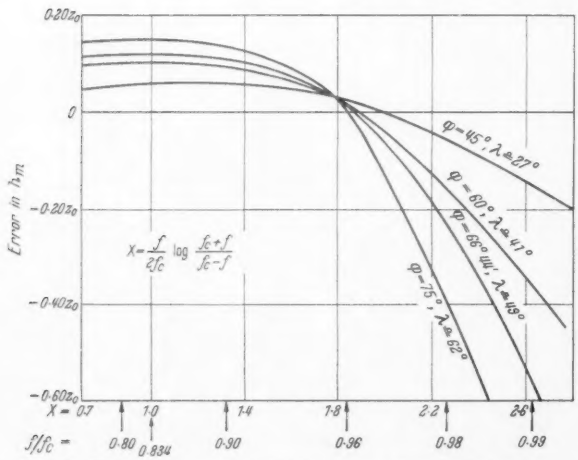


Fig. 8. Error in apparent height of maximum ionisation in a parabolic layer. Critical frequency = 7 Mc/sec.

Thus the percentage error in the apparent z_o is $100\,d\,Y/d\,X$; this is plotted in Figures 7 and 10. The error is always positive, which shows that the apparent thickness of a parabolic layer [estimated from equation (6)] is always greater than its real thickness. For example, an estimate of layer thickness using the slope of an observed (h', f) curve at 96% of the critical frequency will be high by 41% for a critical frequency of 7 Mc/sec at a height of 300 km above England, or 55% for a critical frequency of 3·5 Mc/sec. It is usual in estimations from an observed (h', f) curve to use values of f greater than $0{\cdot}9f_c$, because the behaviour of the ionosphere near the penetration frequency is most important for the prediction of maximum usable frequencies (Appleton and Beynon [9], [10]), and because the group retardation of lower frequencies is often affected appreciably by lower layers in the atmosphere. Therefore practical estimates of layer thickness are usually considerably higher than the true value.

The error in the apparent height of maximum ionisation is given by $z_o \{Y-(X-1)\,dY/dX\}$. This quantity, divided by z_o, is plotted in Figures 8 and 11. It is positive when f/f_c is less than about 0·96, which means that estimates of h_m using this part of the (h', f) curve will be slightly higher than the true value; it is negative when f/f_c is greater than about 0·96, and zero when f/f_c is about 0·96. Since, in estimating h_m from an observed (h', f) curve, only those values of f between $0{\cdot}9 f_c$

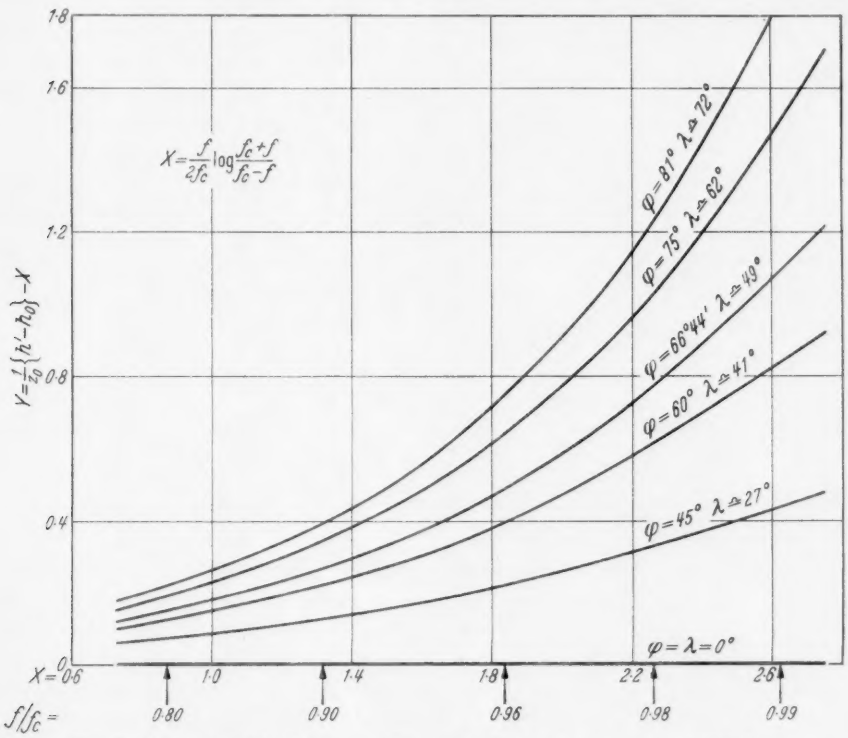


Fig. 9. Computed (h', X) curves. Data as in Figure 6, except $f_c = 3{\cdot}5$ Mc/sec.

and $1{\cdot}0 f_c$ are used, we may surmise that such estimates of h_m are in error by only a small amount; it will be assumed in Section 3 that this error is zero.

The effect of the magnetic field on the group path through a layer may be sensitive to the shape of the layer; calculations similar to the above for, say, a CHAPMAN layer, may yield appreciably different results. However, it would seem worth while to use the above results for the reduction of observed (h', f) curves rather than the no-field theory. This would mean that h' should be plotted not against X but against $(h'-h_o)/z_o$ as given in Tables 2 and 3 (this is equal to $X+Y$ in Figures 6 and 9). Such plots will be straight lines, apart from errors caused by the assumption that the distribution of ionisation is parabolic. In Figure 12 we have plotted data from APPLETON and BEYNONs paper [9] according to this rule. Values from Table 2 (in which it is assumed that $f_c = 7$ Mc/sec at a height of 300 km) have been used for the top two curves, and values from Table 3 (in which it is assumed that $f_c = 3{\cdot}8$ Mc/sec at 120 km) for the bottom curve. The errors caused by the fact that these are not the

exact critical frequencies and heights are probably small; APPLETON and BEYNON do not state to what critical frequencies their curves apply, but they are probably near those which we have assumed. It will be seen that all of the points except two lie very near the appropriate straight line; we have neglected these two points in drawing the middle line; the displacement of the points may be due to retardation in a lower layer. In Table 4 we present estimates of h_m and y_m from Figure 12 and compare them with the corresponding estimates from APPLETON and BEYNONs curves. Estimates of h_m differ little, but estimates of y_m differ considerably. Even if it is desired to estimate only h_m, the method of Figure 12 is preferable to the method of APPLETON and BEYNON; in the latter method some guesswork is generally required in order to draw the best straight line through the experimental points.

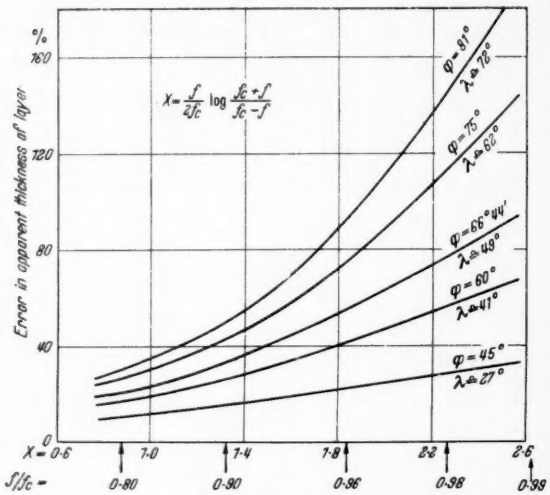


Fig. 10. Error in apparent thickness of a parabolic layer. $f_c = 3{\cdot}5$ Mc/sec.

3. The effect of the magnetic field on oblique incidence propagation

While the deductions of the previous section are of importance when investigating the properties of the atmosphere, it must not be thought that they are necessarily of importance when predicting the behaviour of the ionosphere for the propagation of radio waves at oblique incidence, which is one of the purposes for which vertical incidence ionospheric soundings are made. Calculations of the effect of the magnetic field for oblique incidence are more complicated than those for vertical incidence; consequently we shall only indicate the sort of results to be expected, without coming to detailed numerical conclusions.

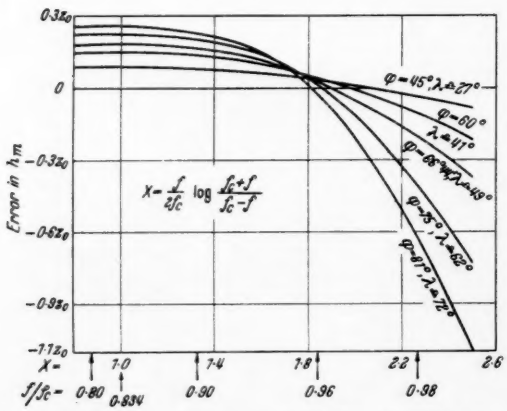


Fig. 11. Error in apparent height of maximum ionisation in a parabolic layer. $f_c = 3{\cdot}5$ Mc/sec.

In order to predict maximum usable frequencies from observations at vertical incidence APPLETON and BEYNON use geometrical considerations in order to determine the angle of incidence of the wave, and MARTYNs Theorem in order to determine how it is propagated in the ionosphere. These together determine a "maximum usable frequency factor," which is defined as that factor by which the critical frequency for vertical incidence at the mid-point of the path must be multiplied in

order to give the maximum usable frequency. We will now discuss corrections which must be applied to the M.U.F. factor in order to take into account the effect of the magnetic field.

One of the consequences of MARTYN Theorem is that, in the absence of a magnetic field, the value of x (proportional to the electronic density) required to reflect a wave, incident on the ionosphere at an angle α to the vertical, is $\cos^2\alpha$. BOOKER, in an important paper on oblique incidence [12], deduces that, in the presence of a magnetic field, for frequencies greater than the gyrofrequency, the value of x required to reflect the ordinary wave lies between $\cos^2\alpha$ and $\cos^2\alpha(1+y)$; it attains the former value for east-west propagation at the magnetic equator, and approaches the latter value for north-south propagation at very oblique incidence near the magnetic equator. He also deduces that the value of x required to reflect the extraordinary wave lies between $\cos^2\alpha(1-y)$ and $1+\cos^2\alpha+\sqrt{(\sin^4\alpha+4y^2\cos^2\alpha)}$. It follows therefore that for frequencies greater than the gyrofrequency, the value of x required for reflection is greater than or equal to $\cos^2\alpha$ for the ordinary wave, and greater than or equal to $(1-y)\cos^2\alpha$ for the extraordinary wave. Thus the critical frequency for a particular angle of incidence is always less than that predicted by MARTYNs Theorem.

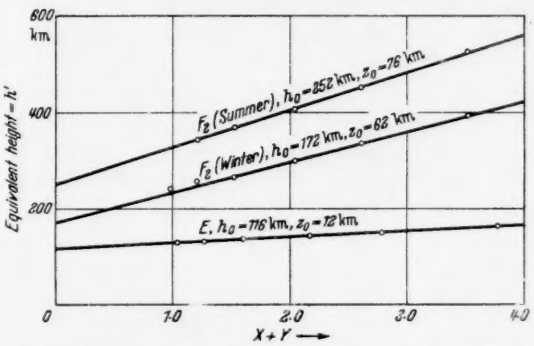


Fig. 12. Suggested method of plotting observations of equivalent height as a function of frequency. The slope of the straight line gives the semi-thickness of the parabolic layer; its intercept on the h' axis gives the height of the lower boundary of the layer. Values of h' taken from APPLETON and BEYNON [9].

MILLINGTON [13] presents values of this critical frequency for an angle of dip of 45°, all values of α (angle of incidence) from 0 to 90°, and NW to SE or NE to SW propagation. For example, let us suppose that $f_H = 0{\cdot}88$ Mc/sec, $f_c = 8{\cdot}8$ Mc/sec, and $\alpha = 45°$. Then the critical frequency at oblique incidence is 17·32 Mc/sec for the ordinary and 18·97 Mc/sec for the extraordinary wave. A prediction of maximum usable frequency by APPLETON and BEYNONs method (*i.e.* MARTYNs Theorem) would give 17·6 Mc/sec for the ordinary wave, and 19·56 Mc/sec for the extraordinary wave; the errors in MARTYNs Theorem are therefore 1·6 % and 3·0 % respectively. For $f_c = 4{\cdot}4$ Mc/sec the errors would be 2·5 % and 6·2 %; the errors are always larger

Table 4. (Note; $y_m = z_0$)

Description of layer	F_2 *(Summer)*	F_2 *(Winter)*	E
Estimates of y_m from the curves in Ref. [9]. . . .	118 km	91 km	22 km
Estimates of y_m from Figure 12	76 km	62 km	12 km
Error in estimate of y_m	58 %	46 %	83 %
Estimates of h_m from the curves in Ref. [9] . . .	328 km	240 km	126 km
Estimates of h_m from Figure 12.	328 km	234 km	128 km

VOL. 2 1952

for a smaller critical frequency. Poeverlein ([16], Figure 10) presents the value of x required to reflect a wave of frequency 3·75 Mc/sec, when $\varphi = 65°$ and $f_H = 1{\cdot}4$ Mc/sec, for N to S propagation and all values of α. Here the discrepancy from Martyns Theorem is about 4% (in frequency) for the ordinary ray, and about $6\frac{1}{2}$% for the extraordinary ray, when $\alpha = 45°$.

The maximum usable frequency for propagation between two points on the earth's surface depends partly on the angle of elevation which the ray must have in order to return to earth at the correct horizontal distance; this depends on the height and thickness of the layer, and on the path of the ray through the layer. Appleton and Beynon ([9], [10]) show that, when the effect of the magnetic field is neglected, the M.U.F. factor is dependent largely on h_m, the height of the maximum of the layer, and almost independent of y_m, the semi-thickness of the layer. For example, an overestimate of y_m by as much as 50% yields a M.U.F. factor which is low by only about 2%. Thus the overestimates of y_m discussed in Section 2 are not of much consequence for the prediction of maximum usable frequencies.

Booker, Millington, and Poeverlein, all give general methods, largely graphical, for computing ray path, time of flight, and attenuation. Several interesting results of this theory are given by Poeverlein ([5], Figures 6 and 7, [16], Table 4) for a frequency of 3·75 Mc/sec in South Germany, and N to S propagation. He finds that rays incident at angles less than 12·7° follow a path with a cusp at the top, and are reflected at the same height as at vertical incidence; also that rays incident at angles near 12·7° are heavily absorbed. He also plots several ray paths, and presents a table which gives height of reflection as a function of horizontal distance for a particular ionospheric layer.

The results presented by these authors are not sufficiently complete to give practical figures for oblique incidence in the presence of a magnetic field similar to those given by Appleton and Beynon neglecting the effect of the field. However, it is to be hoped that a fuller picture will soon be obtained. A different approach, via the Poynting vector, is given by Scott [21]. For the present, no modification to the theory of Appleton and Beynon is suggested except that, for the purpose of estimating layer heights and thicknesses from an observed (h', f) curve, the results shown in Tables 2 and 3, together with interpolated values for other critical frequencies, should be used.

Conclusions

Curves showing the dependence of the vertical component of the group velocity on frequency and electron density are presented, for ordinary and extraordinary waves propagated vertically upwards in South East England. Curves for vertical propagation in any other part of the world will be of similar shape. The effect of the earth's magnetic field is considerable, except for the ordinary wave near the magnetic equator.

The group delays of ordinary and extraordinary waves propagated through a "rectangular" layer are calculated, for several frequencies just greater than the critical frequency of the layer. The results show good agreement with observations of echoes from Region F at frequencies just above the penetration frequency of Region E.

The apparent height of reflection of the ordinary ray is calculated for a "parabolic" layer, for a number of frequencies below its critical frequency. These results are

used to show that the standard method (APPLETON and BEYNON [9]) of estimating the thickness of an ionospheric layer from radio observations yields values which are too high. The amount of this error depends on the magnetic latitude and the critical frequency; curves showing the error as a function of these factors are given in Figures 7 and 10. APPLETON and BEYNONs method, however, yields approximately correct values of the height of the maximum electron density in the ionospheric layer. A modification to their method in order to take into account the effect of the earth's magnetic field is suggested, and illustrated in Figure 12. This Figure shows that the agreement of observation with theory is better when the effect of the magnetic field is taken into account (*i.e.* the experimental points, plotted according to the modified method, lie more nearly on a straight line).

The effect of the magnetic field on oblique incidence propagation is discussed in the light of numerical results obtained in this paper and elsewhere. The picture of oblique incidence propagation is still very incomplete, and, until further results are available, no modification to the existing method of predicting maximum usable frequencies (APPLETON and BEYNON [9], [10]) is suggested. It is shown that the errors in the estimates of layer thickness, although they are considerable, do not cause much error in the maximum usable frequency. The critical frequency for oblique incidence is always less than that predicted by MARTYNs Theorem, but not much less in the cases considered here; the investigation is, however, very incomplete, and the error may be considerable in some cases, *e.g.* near a magnetic pole. The effect of the magnetic field on the absorption of the wave is probably more important at oblique incidence than its effect on the maximum usable frequency. No calculations are performed on this, absorption being neglected throughout the paper. In spite of the fact that the overestimates of layer thickness do not cause much error in estimates of the maximum usable frequency, it is better to use the modified method (Figure 12) of estimating layer thickness and height of maximum ionisation, both because this method should make the reduction of vertical incidence data easier, and because the values of layer thickness may be required for other purposes.

A few values of the apparent height of reflection of the ordinary ray for a "linear" layer are calculated in the Appendix. It is found that, in accordance with results previously obtained (GOUBAU [3], POEVERLEIN [5]), the magnetic field always causes an increase in this height. The increase is greater for a parabolic layer than for a linear layer; the discrepancy between these is very marked near the critical frequency of the parabolic layer.

VOL. 2 1952

APPENDIX

1. *Triple splitting*

The extraordinary wave cannot be propagated in a region where x lies between $(1-y)$ and $(1-y^2)/(1-y^2\cos^2\Theta)$ (assuming that y is less than unity). But it can be propagated in the region where x lies between $(1-y^2)/(1-y^2\cos^2\Theta)$ and $(1+y)$. In this region it is known as the "z-mode." At first sight it would seem that the second region is of no importance, since the extraordinary wave is reflected where $x=(1-y)$, and so cannot penetrate into the upper region. However, the ordinary wave penetrates to the level where $x=1$, and under some circumstances (RYDBECK [14]) it is not totally reflected at that level, but is partly transformed into the z-mode. It is then propagated up to the level where $x=(1+y)$, where it is reflected. It follows the same course on its downward journey. The extraordinary, ordinary, and ordinary-z-ordinary waves are then returned from the ionosphere as three separate echoes. The third echo is called the z-trace.

The coupling between the ordinary ray and the z-ray is strong when Θ is small. It is complete when Θ is zero (at the magnetic poles), when the ordinary ray is not reflected at all, and we are left with the extraordinary ray and the z-ray. Although triple splitting is seldom observed in England, we consider it worth while to present, in Figures 13 and 14, curves which show the behaviour of the group velocity of the extraordinary wave for $y<1$ and values of x up to $(1+y)$.

The curves are similar to those in Figure 5 in that the group velocity is greater near the critical electron density than in the absence of a magnetic field for a wave reflected at the same height. The curves are in two parts, the parts for $0<x<(1-y)$ being the same as in Figure 3, while the other parts of each curve extend from $x/(1+y) = (1-y)/(1-y^2\cos^2\Theta)$ to $x/(1+y) = 1$. In order to give a more general picture of the way in which the shape changes with the angle Θ, Figure 14 shows the curves at a fixed frequency ($y=\frac{1}{2}$) for transverse ($\Theta=90°$), longitudinal, and intermediate propagation.

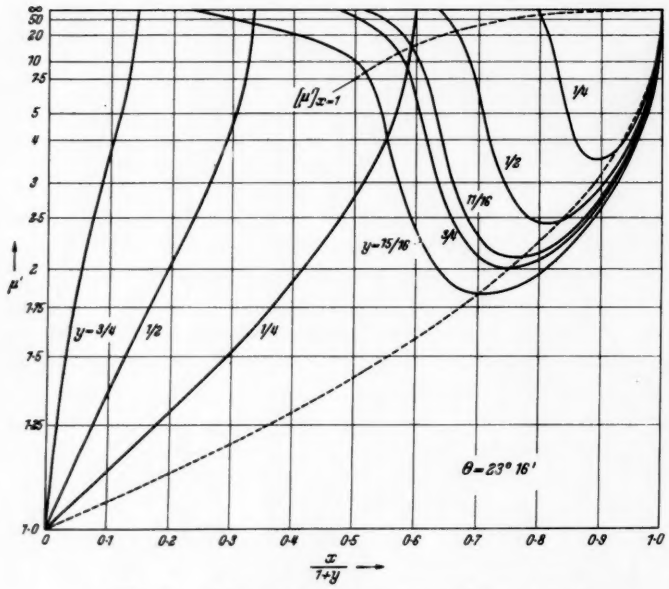


Fig. 13. μ' for extraordinary ray and z-ray, $y<1$.

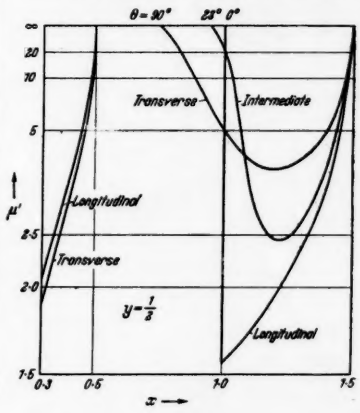


Fig. 14. μ' for extraordinary ray and z-ray, $y<1$.

2. *Another form of (μ', x) curve)*

In a recent paper POEVERLEIN [5] calculates the group delay for vertical and oblique incidence on a linear layer. In this calculation $\mu' t$ is plotted against t, where $t=\sqrt{(1-x)}$. The group delay in a linear layer is proportional to $\int_0^1 \mu'\,dx$ or $2\int_0^1 \mu' t\,dt$. It may therefore be found by measuring the area under either the (μ', x) or the ($\mu' t$, t) curve; the latter is much easier to do because the ordinate is everywhere finite. The device of plotting $\mu'\sqrt{(1-x)}$ against $\sqrt{(1-x)}$ is useful not only for integration (it is exactly the same transformation as we have used for integration through a parabolic layer), but also for displaying μ' as a function of x. The advantages of the method are:

(i) the curve for no magnetic field is a straight line; deviations from this are more readily seen than in Figure 1;

(ii) the method presents directly the relative contributions of different parts of the curve to the integral through a linear layer; modifications for another shape of layer are easily applied;

(iii) the part of the curve of most interest, that approaching $x=1$, is shown on an expanded scale. Curves of this sort are also given by MILLINGTON [4]. A similar transformation exists for the extraordinary wave.

Figure 15 shows Figure 1 redrawn according to this method. Instead of the curve for $y=4$ we have drawn the curve for $y=\infty$. This curve is of no intrinsic interest, since the ray theory is inapplicable when y is large. However, it is interesting in showing a limit within which the other

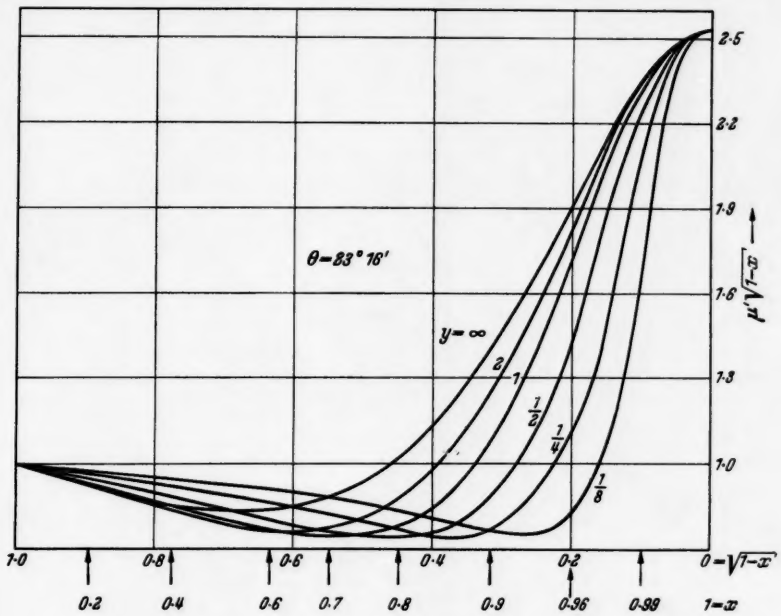


Fig. 15. $\mu'\sqrt{(1-x)}$ for ordinary ray $\left(=\dfrac{\mu'}{\mu' \text{ for no field}}\right)$.

curves lie. For x greater than about 0·5 all curves for $y>2$ lie between the curve for $y=2$ and that for $y=\infty$. For x less than about 0·5 the curves are nearly the same for all y greater than 2. The curve for $y=\infty$ was computed from the formulae

$$\mu^2 = \frac{(1-x)\operatorname{cosec}^2\Theta}{1+(1-x)\cot^2\Theta}$$

and

$$\mu\mu'-1 = \frac{x^2\mu^4\sin^2\Theta\cos^2\Theta}{(1-x)^2}\,.$$

The areas under the curves in Figure 15 give directly the apparent heights of reflection for a linear layer above England. These heights are shown as a function of frequency in Table 5. The ordinary ray travels more quickly in the lower part of its path, and more slowly in the upper part of its path, than it would in the absence of a magnetic field. It is interesting to notice, from Table 5, that the net effect of the magnetic field is to cause a delay, *i.e.* to increase the apparent height. This is in accordance with the results of Goubau [3] and Poeverlein [5].

Table 5. Increase in apparent height, h_o, due to the presence of a magnetic field. Ordinary ray. Linear layer. In England

Frequency / Gyrofrequency	0	$\frac{1}{2}$	1	2	4	8
Increase in apparent height	30·0 %	23·3 %	18·7 %	13·8 %	8·9 %	5·1 %

VOL. 2 1952

3. *A note on accuracy*

The only errors to be expected in the evaluation of the group velocity are rounding off errors; the computation was so arranged that these were negligible. But errors of some importance may

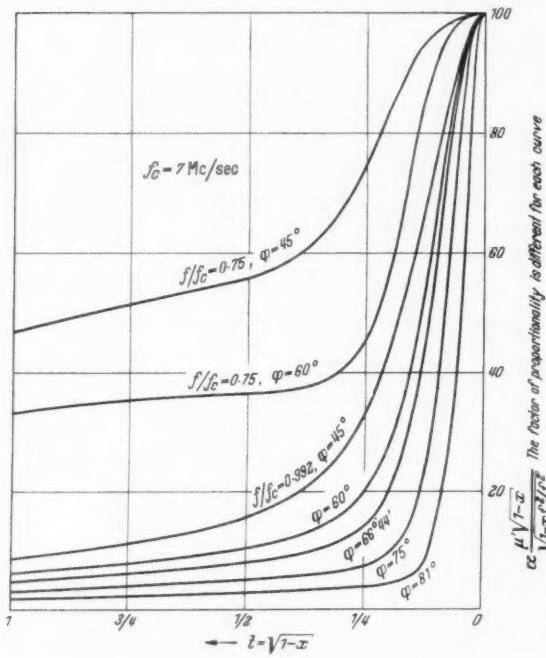


Fig. 16a.

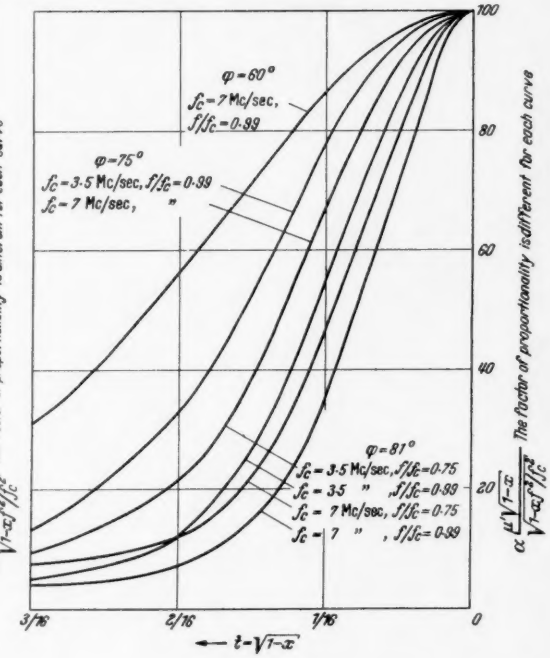


Fig. 16b.

Fig. 16a. The integrand in equation (5) for $f_c = 7$ Mc/sec and various values of φ and f/f_c.

Fig. 16b. The integrand in equation (5) over part of its range. The ordinates at $t = 0$, 1/16, 2/16, and 3/16, are shown, indicating the inaccuracy of integration in this range using only these four ordinates.

be introduced by the inaccuracy of SIMPSONS Rule. An appreciable error is caused when the integrand varies considerably between successive ordinates. This happens when φ is large (at high latitudes); the error increases as f/f_c increases, and is larger for f_c=7 Mc/sec than for f_c=3·5 Mc/sec.

Table 6. Estimated error in numerical integration

φ	f_c (Mc/sec)	*Error when* $f/f_c = 127/128$ %	*Error when* $f/f_c = 3/4$ %
81°	3·5	±3	±1½
75°	7	±1·8	±0·6
75°	3·5	±0·8	±0·2
66° 44′	7	±0·4	±0·1
66° 44′	3·5	±0·2	±0·05
0°	any value	−0·021	<0·001

Figure 16a shows curves of the integrand in equation (5) for several values of the parameters φ, f_c, and f/f_c. The sharp rise near the end of the curves when φ is large, makes the evaluation of the integral using only 17 ordinates inaccurate. The last four ordinates are shown on an enlarged scale in Figure 16b. Of the curves shown in this figure, the lowest two (for φ=81°, and f_c=7 Mc/sec) yield results so inaccurate that these results have been discarded; this explains the absence of results in the last column of Table 2. The estimated errors in the results from the other curves are shown in Table 6. It will be seen from this table that the results for φ=81°, f_c=3·5 Mc/sec, should be treated with reserve, but the other results are within

the limit of normal experimental accuracy. It might be worth while to repeat some of these calculations using a closer spacing of ordinates in the range $0<t<1/4$.

The error caused by using SIMPSON's Rule was found exactly for $\varphi=0°$; in this case the integral can be expressed analytically by equation (6). The errors in the other integrals were obtained partly by guessing the possible shapes of the curves in the range $0<t<1/4$, and partly from a comparison of SIMPSON's and WEDDLE's Rules in the range $0<t<3/8$.

Acknowledgements—The computations were carried out on the EDSAC at the University Mathematical Laboratory, Cambridge, through the courtesy of the Director, Dr. M. V. WILKES. During most of this time we were both members of the Radio Section of the Cavendish Laboratory, Cambridge. The Radio Section was financed by a grant from the Department of Scientific and Industrial Research. D. H. SHINN was supported by a grant from D.S.I.R. and H. A. WHALE by an 1851 Science Research Scholarship and a Rouse Ball Research Studentship from Trinity College, Cambridge. The work was completed by D. H. SHINN at Radio Research Station, Slough, and, later, at Marconi's Wireless Telegraph Company. We wish to express our appreciation of the support received from these sources.

We are greatly indebted to our colleagues in the Cavendish Laboratory and the University Mathematical Laboratory, to Mr. W. R. PIGGOTT of Radio Research Station, and to Mr. G. MILLINGTON, of Marconi's Wireless Telegraph Company, for the numerous valuable discussions which we have had with them.

VOL. 2 1952

REFERENCES

[1] APPLETON, E. V.; J. Inst. Elect. Engrs 1932 **71** 642. [2] BAJPAI, R. R. and MATHUR, K. B.; Indian J. Phys. 1937 **11** 165. [3] GOUBAU, G.; Hochfrequenz u. Electroakust. 1934 **44** 17. [4] MILLINGTON, G.; Proc. Phys. Soc. 1938 **50** 561. [5] POEVERLEIN, H.; Z. angew. Phys. 1951 **3** 135. [6] DE VOOGT, A. H.; Tijds. Nederland. Radiogenootschap 1948 (Nov) 183. [7] WHALE, H. A. and STANLEY, J. P.; J. Atmosph. Terr. Phys. 1950 **1** 82. [8] APPLETON, E. V.; Proc. Roy. Soc. A 1937 **162** 451. [9] APPLETON, E. V. and BEYNON, W. J. G.; Proc. Phys. Soc. 1940 **52** 518. [10] APPLETON, E. V. and BEYNON, W. J. G.; Proc. Phys. Soc. 1947 **59** 58. [11] BOOKER, H. G. and SEATON, S. L.; Phys. Rev. 1940 **57** 87. [12] BOOKER, H. G.; J. Geophys. Res. 1949 **54** 243. [13] MILLINGTON, G.; Inst. Elect. Engrs Monograph No. 1, 1951 (April). [14] RYDBECK, O. E. H.; J. Appl. Phys. 1950 **21** 1205. [15] POEVERLEIN, H.; Z. angew. Phys. 1949 **1** 517. [16] POEVERLEIN, H.; Z. angew. Phys. 1950 **2** 152. [17] RATCLIFFE, J. A.; Wireless Eng. 1933 **10** 354. [18] RYDBECK, O. E. H.; Phil. Mag. 1940 **30** 282. [19] WHALE, H. A.; J. Atmosph. Terr. Phys. 1951 **1** 244. [20] BOOKER, H. G.; "Studies on propagation in the ionosphere; Technical Report No. 1," School of Electrical Engineering, Cornell University. [21] SCOTT, J. W. C.; Proc. Inst. Radio Engrs 1950 **38** 1057.

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 106 to 129. Pergamon Press Ltd., London

The anomalous propagation of radio waves in the 1–10 metre band

F. H. NORTHOVER

Dundee Technical College

(Received 8 August 1951)

ABSTRACT

Anomalous propagation of metre waves, which is occasionally observed, cannot be explained as in the case of centimetre waves by the effects of surface inversions, but L. E. BEGHIAN has suggested that it may be caused by sharp inversions occurring at high levels. The theory developed in this paper shows that high level subsidence inversions, which sometimes occur, can in fact account satisfactorily for the anomalous propagation in the 1–10 metre band.

1. INTRODUCTION

1.1. It is well known that anomalous propagation of V.H.F. (very high frequency) radio waves is due to local upward decrements of refractive index μ caused by temperature inversions and/or upward lapses in vapour pressure. These fall into two types conveniently designated as "surface inversions" and high "level inversions."

In surface inversions, there is a steady decrease in refractive index μ with height throughout a region near the earth's surface. Their effect on propagation has been studied by many writers, notably BOOKER [1], and PEKERIS and AMENT [2]. They cause anomalous propagation on the shorter V.H.F. waves only (*i. e.* around Britain, on waves shorter than about 30 cm).

High level subsidence inversions were first systematically described by WESTWATER [3] although previously noticed in experiments at San Diego [4]. They are due to a mass of warm dry air overlying a cool moist air mass, and usually occur at heights of about 1000 m during settled anticyclonic weather. Such conditions often cause sharp lapses in μ, and it was suggested by L. E. BEGHIAN in 1944, that these inversions might be responsible for the anomalous propagation of 3·5 m waves observed in some Admiralty trials carried out in 1943 between Brixham, Devon and St. Catherine's Point, Isle of Wight.

1.2. *The experimental evidence*

BEGHIANs suggestion, and some preliminary calculations of inversion layer reflection coefficients made by him and the writer [5] led to the systematic Admiralty trials mentioned by MEGAW [6]. These were carried out in 1945 across a 100-mile land link between Coventry and Chelmsford on 3·5 m at the same time as the General Electric Company's trials on 3 and 9 cm in Cardigan Bay. The transmitter power was 100 w, radiated from a half-wave directional horizontal array at a height of 30 m. The height of the receiving antenna was 5 m.

Upper air measurements were taken with specially sensitive apparatus and a high level inversion was discovered, details of which are given by WESTWATER [3]. See also SHEPPARD [7].

Metre wave propagation was observed only when this high level inversion was present and was not due to surface inversions, since it continued when these were absent.

Since the war, further experimental work, carried out by SAXTON [8] of the National Physical Laboratory, has confirmed the importance of high level inversions in metre wave propagation. Using reflection coefficient formulae developed by BEGHIAN and the writer [5], he has made some interesting field strength calculations covering ranges lying well within the greatest distance (hereinafter referred to as the "limiting distance") which can be reached by a ray reflected only *once* from the inversion. Unfortunately, his methods cannot be used at distances near to, or exceeding, the limiting distance since they do not take account of diffraction around the earth's surface which then becomes important. The present paper attempts to provide a method for calculating field strengths at such distances by placing the theory upon a rigorous mathematical basis. The convergence of the field series developed, while usually sufficiently rapid to yield detailed information about the field near to, or beyond the limiting distance, is too slow for it to be useful much within this distance. The latter case has, however, already been discussed by SAXTON [9].

2. *Reflection from a high level inversion*

VOL. 2 1952

2.1. At ranges near to, and beyond, the limiting distance, it will be shown that the field strength can be expressed as a series whose convergence is usually sufficient for the first few terms to give a good approximation to its sum. It appears that each term in the series can be interpreted as a ray leaving the transmitter and incident upon the layer at an obliquity which decreases as its order in the series increases. The dominant terms at the beginning correspond to rays which leave the transmitter and strike the layer at nearly "limiting"* obliquity. Furthermore, the attenuation of each "mode" is directly expressible in terms of the reflection coefficient, at the layer, of the ray which it represents.

There is thus a close connection between wave theory and elementary ray theory. We therefore begin by investigating the reflection of plane waves from an elevated inversion.

We take axes Oy in the lower face of the inversion layer, and Oz vertically. The components of the field vectors then satisfy the equation,

$$\frac{\partial^2 X}{\partial y^2} + \frac{\partial^2 X}{\partial z^2} + \frac{p^2}{C^2} \varkappa X = 0 \tag{1}$$

where $p/2\pi$ is the frequency, C the velocity of light in vacuo, and $\varkappa$ is the dielectric constant at any point within the layer.

Let the incident and transmitted waves make angles χ, Θ respectively with the layer; then, if $\varkappa_0$ and unity are respectively the dielectric constants below and above the layer, the incident, reflected, and transmitted waves are respectively represented by the expressions

$$\operatorname{Exp}\left\{\frac{ip}{C}\sqrt{\varkappa_0}\,(y\cos\chi + z\sin\chi)\right\}, \qquad R\operatorname{Exp}\left\{\frac{ip}{C}\sqrt{\varkappa_0}\,(y\cos\chi - z\sin\chi)\right\}$$

* The obliquity, at the layer, of the ray defining the "limiting distance."

and

$$R' \operatorname{Exp}\left\{\frac{ip}{C}(y\cos\Theta + z\sin\Theta)\right\},$$

where R and R' are respectively the complex reflection and transmission coefficients.

Within the layer, we write

$$X = X_0 \operatorname{Exp}\left(\frac{ip}{C} y\cos\Theta\right).f(z) \tag{2}$$

then $f(z)$ satisfies the equation

$$f''(z) + \frac{p^2}{C^2}(\varkappa - \cos^2\Theta) f(z) = 0. \tag{3}$$

Put

$$z = \tau\gamma \tag{3A}$$

where τ is the layer thickness; then $f(\gamma)$ satisfies,

$$f''(\gamma) + \vartheta^2(\varkappa - \cos^2\Theta).f(\gamma) = 0 \tag{4}$$

where

$$\vartheta = 2\pi\tau/\lambda. \tag{4A}$$

The boundary conditions require, first, that

$$\sqrt{\varkappa_0}.\cos\chi = \cos\Theta \tag{5}$$

so that equation (4) becomes,

$$f''(\gamma) + \vartheta^2(\varkappa - \varkappa_0\cos^2\chi) f(\gamma) = 0. \tag{6}$$

In applications to high level inversions, χ is small, and $\varkappa$ is nearly unity. Hence, writing,

$$\varkappa = 1 + \delta, \qquad \varkappa_0 = 1 + \delta_0 \qquad (\delta, \delta_0 \ll 1) \tag{7}$$

equation (6) becomes,

$$f''(\gamma) + \vartheta^2(\delta - \delta_0 + \chi^2) f(\gamma) = 0 \tag{8}$$

approximately.

If $u(\gamma)$, $v(\gamma)$ are a pair of fundamental solutions of (6), the boundary conditions at the layer are,

$$\left.\begin{aligned} 1 + R &= A\,u(0) + B\,v(0) \\ R' \operatorname{Exp}\left(\frac{ip\tau}{C}\sin\Theta\right) &= A\,u(1) + B\,v(1) \\ (1-R)\frac{ip\tau}{C}\sqrt{\varkappa_0}.\sin\chi &= A\,u'(0) + B\,v'(0) \\ R' \operatorname{Exp}\left(\frac{ip}{C}\tau\sin\Theta\right).\frac{ip\tau}{C}\sin\Theta &= A\,u'(1) + B\,v'(1). \end{aligned}\right\} \tag{9}$$

These equations give

$$R = \frac{(L\sqrt{\varkappa_0}\sin\chi\sin\Theta - N) + i(M_1\sqrt{\varkappa_0}\sin\chi + M_2\sin\Theta)}{(L\sqrt{\varkappa_0}\sin\chi\sin\Theta + N) + i(M_1\sqrt{\varkappa_0}\sin\chi - M_2\sin\Theta)} \tag{10}$$

where

$$\left.\begin{aligned} L &= u(1)\,v(0) - u(0)\,v(1) \\ \vartheta\,M_1 &= u'(1)\,v(0) - u(0)\,v'(1) \\ \vartheta\,M_2 &= u(1)\,v'(0) - u'(0)\,v(1) \\ \vartheta^2 N &= u'(1)\,v'(0) - u'(0)\,v'(1). \end{aligned}\right\} \tag{11}$$

For high level inversions χ and δ_0 are small; hence R approximates closely to the value

$$\frac{\{L\chi(\chi^2-\delta_0)^{\frac{1}{2}}-N\}+i\{M_1\chi+M_2(\chi^2-\delta_0)^{\frac{1}{2}}\}}{\{L\chi(\chi^2-\delta_0)^{\frac{1}{2}}+N\}+i\{M_1\chi-M_2(\chi^2-\delta_0)^{\frac{1}{2}}\}}. \tag{12}$$

This tends to the Fresnel value $\{\chi-(\chi^2-\delta_0)^{\frac{1}{2}}\}/\{\chi+(\chi^2-\delta_0)^{\frac{1}{2}}\}$ as $\tau\to 0$, since it is easy to shew, from (8), that, as $\vartheta\to 0$; $L\to 1$, $M_2\to\vartheta^{-1}\to -M_1$, and $N\to\chi^2$.

2.2. *The inversion profile*

Before further progress can be made, it is necessary to know how the refractive index varies through the inversion layer. Unfortunately, very few ascents with the special apparatus described by WESTWATER and SHEPPARD seem to have been made. In fact the only record of a typical high level inversion which is readily available seems to be that given in SHEPPARDs paper. It will repay very careful study. Details of this inversion (hereafter referred to as the "Cardington inversion"), are,

	Dry	Wet
Temperature below inversion (° F)	41	40
Temperature above inversion	51	40

Height of inversion base $= 3900$ ft $\doteqdot 1200$ m. The decrease in dry bulb temperature during the first 20 ft was 5° F, and the pressure at the inversion base about 1000 mb.

Using the formula

$$\mu-1=\frac{79}{T}\left(p-\frac{e}{7}+\frac{4800\,e}{T}\right).10^{-6} \tag{13}$$

where T is the absolute temperature, p the pressure in millibars, and e the vapour pressure, we find that the total change in μ, *viz.* $(\Delta\mu)_0$ say, is $2{\cdot}58\times10^{-5}$, and the change in the first 20 ft is $1{\cdot}33\times10^{-5}$. In the opinion of the present writer, the parabolic law

$$\Delta\mu=(\Delta\mu)_0\,(1-\gamma)^2 \tag{14}$$

is the simplest law which will give a satisfactory fit in this instance.

Owing to present lack of knowledge of the detailed structure of high level inversions it will not be unreasonable to take this inversion as typical. In this paper, therefore, we shall consider only the parabolic law (14). Writing, therefore,

$$\delta=\delta_0\,(1-\gamma)^2 \tag{15}$$

equation (8) becomes

$$f''(\gamma)+(A+C\gamma+B\gamma^2)f(\gamma)=0 \tag{16}$$

where

$$A=\vartheta^2\chi^2,\qquad B=\vartheta^2\,\delta_0,\qquad C=-\,2\,\delta_0\vartheta^2. \tag{16A}$$

In practice, δ_0/χ^2 is always small, and so B/A, C/A are small.

2.3. *Thin layers.* $(\vartheta\chi\leq 1)$

Since B/A, C/A are small we may, in approximating solutions of (16) as far as 2nd order small quantities, include A^2, AB, AC; but neglect B^2, BC, C^2 and 3rd order

VOL. 2 1952

terms. The approximate solutions of (16) then are,

$$\left.\begin{aligned} u(\gamma) &= 1-\frac{1}{2}A\gamma^2-\frac{1}{6}C\gamma^3-\frac{1}{12}B\gamma^4+\frac{1}{24}A^2\gamma^4+\frac{1}{30}AC\gamma^5+\frac{7}{360}AB\gamma^6 \\ v(\gamma) &= \gamma-\frac{1}{6}A\gamma^3-\frac{1}{12}C\gamma^4-\frac{1}{20}B\gamma^5+\frac{1}{120}A^2\gamma^5+\frac{1}{120}AC\gamma^6+\frac{13}{2520}AB\gamma^7. \end{aligned}\right\} \quad (17)$$

Using these values of $u(\gamma)$ and $v(\gamma)$ to work out L, M_1, M_2, and N and remembering that $\delta_0/\chi^2 \ll 1$, we find, from (12), that

$$R=\frac{\delta_0}{4\chi^2}\,\frac{i\left(1-\frac{1}{6}\vartheta^2\chi^2+\frac{1}{120}\vartheta^4\chi^4\right)+\frac{1}{3}\vartheta\chi\left(1-\frac{1}{10}\vartheta^2\chi^2\right)}{-i\left(1-\frac{1}{2}\vartheta^2\chi^2+\frac{1}{24}\vartheta^4\chi^4\right)-\left(1-\frac{1}{6}\vartheta^2\chi^2\right)\vartheta\chi} \quad (18)$$

giving,

$$|R|\doteqdot\frac{\delta_0}{4\chi^2}\left(1-\frac{1}{9}\vartheta^2\chi^2+\frac{2}{405}\vartheta^4\chi^4\right) \quad (19)$$

for $\vartheta\chi \leq 1$, say.

2.4. *Thick layers.* ($\vartheta\chi>1$)

The complex reflection coefficient R can be put into the form

$$R=-\frac{\left[\frac{u'(0)}{u(0)}-i\vartheta\chi\right]\left[\frac{v'(1)}{v(1)}-i\vartheta\chi\left(1-\frac{\delta_0}{\chi^2}\right)^{\frac{1}{2}}\right]\frac{v(1)u(0)}{v(0)u(1)}-\left[\frac{v'(0)}{v(0)}-i\vartheta\chi\right]\left[\frac{u'(1)}{u(1)}-i\vartheta\chi\left(1-\frac{\delta_0}{\chi^2}\right)^{\frac{1}{2}}\right]}{\left[\frac{u'(0)}{u(0)}+i\vartheta\chi\right]\left[\frac{v'(1)}{v(1)}-i\vartheta\chi\left(1-\frac{\delta_0}{\chi^2}\right)^{\frac{1}{2}}\right]\frac{v(1)u(0)}{v(0)u(1)}-\left[\frac{v'(0)}{v(0)}+i\vartheta\chi\right]\left[\frac{u'(1)}{u(1)}-i\vartheta\chi\left(1-\frac{\delta_0}{\chi^2}\right)^{\frac{1}{2}}\right]}. \quad (20)$$

If $\vartheta\chi>1$, it may be shown that JEFFREYs approximation (the W.K.B. type of approximation) may be used in (16); hence, approximately,*

$$u(\gamma)=\operatorname{Exp}\{-iF(\gamma)\} \qquad v(\gamma)=\overline{u(\gamma)} \quad (21)$$

where

$$F(\gamma)=\int_0^\gamma\left\{\vartheta(\chi^2+\delta-\delta_0)^{\frac{1}{2}}-\frac{i\delta_0\Phi'(\gamma)}{4(\chi^2+\delta-\delta_0)}-\frac{\delta_0\Phi''(\gamma)}{8\vartheta(\chi^2+\delta-\delta_0)^{\frac{3}{2}}}\right\}d\gamma \quad (21\text{A})$$

$$\left.\begin{aligned} u'(1)/u(1) &= \left\{-i\vartheta(\chi^2-\delta_0)^{\frac{1}{2}}+\frac{i\delta_0\Phi''(1)}{8\vartheta(\chi^2-\delta_0)^{\frac{3}{2}}}\right\} \\ u'(0)/u(0) &= \left\{-i\vartheta\chi-\frac{\delta_0\Phi'(0)}{4\chi^2}+\frac{i\delta_0\Phi''(0)}{8\vartheta\chi^3}\right\}. \end{aligned}\right\} \quad (22)$$

After some reduction we find from (20), that

$$|R|=\frac{\delta_0}{4\chi^2}\cdot\frac{|\Phi'(0)|}{2\vartheta\chi}\{1+(f/\vartheta\chi)\sin[2F(1)]+(f/\vartheta\chi)^2\sin^2[F(1)]\}^{\frac{1}{2}} \quad (23)$$

where

$$f=\Phi''(1)/\Phi'(0) \quad (23\text{A})$$

$$\doteqdot\frac{\delta_0}{4\chi^2}\,\frac{|\Phi'(0)|}{2\vartheta\chi}\left\{1+\frac{f}{\vartheta\chi}\sin 2\vartheta\chi+\left(\frac{f}{\vartheta\chi}\right)^2\sin^2\vartheta\chi\right\}^{\frac{1}{2}} \quad (24)$$

since by (21A),

$$F(1)=\vartheta\chi+0\left(\frac{\delta_0}{2\chi^2}\right).$$

For the parabolic layer $\Phi''(1)=2$, $\Phi'(0)=-2$; hence, when $\vartheta\chi>1$

$$|R|\doteqdot\left(\frac{\delta_0}{4\chi^2}\right)\frac{1}{\vartheta\chi}\left\{1-\frac{1}{\vartheta\chi}\sin 2\vartheta\chi+\frac{1}{\vartheta^2\chi^2}\sin^2\vartheta\chi\right\}^{\frac{1}{2}}. \quad (25)$$

* These are different independent solutions from (17).

Note in (19) and (25), that $(\delta_0/4\chi^2)$ is the Fresnel discontinuity value of $|R|$; also that these two formulae give excellent agreement when $\vartheta\chi = 1$. It can now be seen that the value of $|R|$ given by (19) and (25) agrees closely with that given by earlier profiles ([5], [9] 1951), when $\vartheta\chi <$ about π, but is more favourable when $\vartheta\chi$ is larger; *e.g.* if $\vartheta\chi = 2\pi$, the exponential profile of BEGHIAN gives $|R| = \frac{1}{10} R_F$; whereas, by (25), $|R|$ is then about $\frac{1}{6} R_F$. (R_F is the Fresnel value of $|R|$.)

3. *The field beneath an elevated inversion*

3.1. Let the field be due to a horizontal dipole situated below the inversion at height H above the ground, and be emitting waves of frequency $p/2\pi$. The atmosphere below and above the inversion is supposed homogeneous, and the dielectric constant decreases from $\varkappa_0$ below it to unity above, through a small height interval τ. In discussing the propagation it is not possible [10] to employ the "earth flattening" technique commonly used in dealing with surface inversions owing to the great height at which the changes in $\mu (= \sqrt{\varkappa})$ occur, and their extreme sharpness.

We take spherical polar coordinates (r, Θ, φ) with origin at the earth's centre, initial line $(\Theta = 0)$ along the line joining the earth's centre to the transmitter, and the direction of the dipole axis parallel to $\varphi = \pi/2$. It is shown in the Appendix that, for points whose distance from the dipole is large compared with the wavelength λ and the inversion height h, but small compared with the earth's radius R, the "undisturbed" field* in the direction of optimum propagation is given approximately by**

$$\left.\begin{aligned} \mathsf{E} &= -\frac{ip}{C\varkappa_0} \operatorname{curl} \Lambda_0 \\ \mathsf{H} &= \operatorname{curl}\left(\frac{1}{\varkappa_0} \operatorname{curl} \Lambda_0\right) \end{aligned}\right\} \tag{26}$$

where Λ_0 is the radial vector $(r\varkappa_0 \Pi_0, 0, 0)$

$$\Pi_0 = \frac{\beta}{R_1} \exp\left(-\frac{ip\sqrt{\varkappa_0} R_1}{C}\right)$$

$$R_1^2 = r^2 + (R+H)^2 - 2r(R+H)\cos\Theta .$$

C is the velocity of light in vacuo, and

$$\beta = \frac{\lambda^2}{4\pi^2 R}\left(\frac{3\Gamma}{C\varkappa_0^3}\right)^{\frac{1}{2}}$$

where Γ, the radiated power, is equal to the mean value of $\frac{c}{4\pi}\iint \mathsf{E}_\Lambda \mathsf{H} . d\mathsf{S}$, taken over a large closed surface surrounding the dipole, through one oscillation.

The components of E and H for the undisturbed field are, from (26)

$$E_r = 0, \qquad E_\Theta = 0, \qquad E_\Phi = \frac{ip}{C}\frac{\partial \Pi_0}{\partial \Theta}$$

$$H_r = \frac{ip}{C}\frac{1}{r\sin\Theta}\frac{\partial}{\partial\Theta}\left(\sin\Theta\frac{\partial\Pi_0}{\partial\Theta}\right), \qquad H_\Theta = -\frac{ip}{C}\frac{\partial}{r\,\partial r}\left(r\frac{\partial\Pi_0}{\partial\Theta}\right), \qquad H_\Phi = 0. \tag{26A}$$

* *i.e.* field due to dipole alone.

** We take field vectors proportional to $\exp(ipt)$.

VOL. 2 1952

3.2. *The field inside the inversion*

The appropriate generalizations of (26) when $\varkappa$ varies (within the inversion layer) are,

$$\mathsf{E} = -\frac{ip}{C\varkappa}\,\mathrm{curl}\,\Lambda \tag{27}$$

$$\mathsf{H} = \mathrm{curl}\left(\frac{1}{\varkappa}\,\mathrm{curl}\,\Lambda\right),$$

where Λ is the radial vector

$$\{r\varkappa\Pi(r, \Theta), \qquad 0, 0\}.$$

Using the MAXWELL relation $\mathrm{curl}\,\mathsf{H} = \frac{ip\varkappa}{C}\mathsf{E}$, we find that Π satisfies the equation

$$\frac{1}{r}\frac{\partial^2}{\partial r^2}\left(r\frac{\partial \Pi}{\partial \Theta}\right) + \frac{1}{r^2}\frac{\partial}{\partial \Theta}\left\{\frac{1}{\sin\Theta}\frac{\partial}{\partial \Theta}\left(\sin\Theta\frac{\partial \Pi}{\partial \Theta}\right)\right\} + \frac{p^2}{C^2}\varkappa\frac{\partial \Pi}{\partial \Theta} = 0 \tag{28}$$

and so can be expressed in the form

$$\sum_{(n)}\frac{\Omega_n(r)}{r}P_n(\cos\Theta) \tag{29}$$

where $P_n(\cos\Theta)$ is LEGENDRE's function of order n, and $\Omega_n(r)$ satisfies the equation

$$\frac{d^2\Omega_n}{dr^2} + \left\{\frac{p^2}{C^2}\varkappa - \frac{n(n+1)}{r^2}\right\}\Omega_n = 0 \tag{30}$$

i.e.

$$\frac{d^2\Omega}{d\xi^2} + \left\{\varkappa(\xi) - \frac{n(n+1)}{\xi^2}\right\}\Omega = 0 \tag{31}$$

where

$$\xi = 2\pi r/\lambda. \tag{31A}$$

3.3. *The boundary conditions*

It can be shown that, for general V.H.F. propagation over sea and for metre wave propagation over moist land, the earth's surface can be treated approximately as perfect conductor. Hence by (26A) the approximate condition at the earth's surface is,

$$\frac{\partial \Pi}{\partial \Theta} = 0. \tag{32}$$

Remembering (29), we assume that the disturbed field below the inversion is of the form

$$-\frac{i\beta}{rb}\sum(2n+1)\{A_n\zeta_n(\xi_1) + B_n\psi_n(\xi_1)\}P_n(\cos\Theta) \tag{33}$$

where

$$\zeta_n(\xi_1) = (\tfrac{1}{2}\pi\xi_1)^{\frac{1}{2}}H^{(2)}_{n+\frac{1}{2}}(\xi_1), \qquad \psi_n(\xi_1) = (\tfrac{1}{2}\pi\xi_1)^{\frac{1}{2}}J_{n+\frac{1}{2}}(\xi_1), \qquad \xi_1 = \xi\sqrt{\varkappa_0} \tag{33A}$$

and write the above expression for Π_0 in the form*

$$-\frac{i\beta}{rb}\sum_{n=0}^{\infty}(2n+1)\,\zeta_n(b)\,\psi_n(\xi_1)\,P_n(\cos\Theta) \qquad (\xi_1 < b) \tag{34}$$

where

$$b = 2\pi(R+H)\sqrt{\varkappa_0}/\lambda \tag{34A}$$

* See, for instance, WATSONS paper. If $\xi_1 > b$, ξ_1 and b have to be interchanged.

the surface condition becomes

$$A_n \zeta_n(x_1) + \{B_n + \zeta_n(b)\}\psi_n(x_1) = 0 \tag{35}$$

where

$$x_1 = 2\pi R\sqrt{\varkappa_0}/\lambda. \tag{35A}$$

Let $u_n(\xi)$, $v_n(\xi)$ be any pair of fundamental solutions of (31); then by (29) the form to assume for Π within the inversion is

$$-\frac{i\beta}{rb}\sum (2n+1)\{C_n u_n(\xi) + D_n v_n(\xi)\} P_n(\cos\Theta). \tag{36}$$

Above the inversion, $\varkappa = 1$ and the appropriate form for Π there is

$$-\frac{i\beta}{rb}\sum (2n+1) F_n \zeta_n(\xi) . P_n(\cos\Theta). \tag{37}$$

The boundary conditions at the inversion require that H_θ and E_ϕ should be continuous at its lower and upper edges; *i.e.* by (27), that Π and $\partial(r\Pi)/\partial r$ should be continuous there. This gives

$$\left.\begin{aligned} \{A + \psi(b)\}\zeta(y_1) + B\psi(y_1) &= C u(y_1) + D v(y_1) \\ \{A + \psi(b)\}\zeta'(y_1) + B\psi'(y_1) &= C u'(y_1) + D v'(y_1) \end{aligned}\right\} \tag{38}$$

at the lower edge, and

$$\left.\begin{aligned} C u(y) + D v(y) &= F\zeta(y) \\ C u'(y) + D v'(y) &= F\zeta'(y) \end{aligned}\right\} \tag{39}$$

at the upper edge.

Here

$$y_1 = 2\pi(R+h)\sqrt{\varkappa_0}/\lambda, \qquad y = 2\pi(R+h+\tau)/\lambda \tag{39A}$$

and the suffix n has been suppressed for convenience in writing.

3.4. *The total field below the inversion*

Solving equations (35–39) we obtain,

$$A = -\psi(x_1)\{\zeta(b) - k\eta(b)\}/\{\zeta(x_1) - k\eta(x_1)\}$$
$$B = -k\{\eta(b)\zeta(x_1) - \zeta(b)\eta(x_1)\}/\{\zeta(x_1) - k\eta(x_1)\}$$

where

$$k = \frac{N\zeta(y_1)\zeta(y) + L\zeta'(y_1)\zeta'(y) - M_1\zeta'(y_1)\zeta(y) - M_2\zeta(y_1)\zeta'(y)}{N\eta(y_1)\zeta(y) + L\eta'(y_1)\zeta'(y) - M_1\eta'(y_1)\zeta(y) - M_2\eta(y_1)\zeta'(y)} \tag{40}$$

and L, M_1, M_2 and N are layer functions defined by

$$\left.\begin{aligned} L &= u(y)v(y_1) - u(y_1)v(y) \\ M_1 &= u'(y)v(y_1) - u(y_1)v'(y) \\ M_2 &= u(y)v'(y_1) - u'(y_1)v(y) \\ N &= u'(y)v'(y_1) - u'(y_1)v'(y) \end{aligned}\right\} \tag{40A}$$

VOL. 2 1952

and this gives, for the total field function Π, below the level of the dipole position,

$$\left.\begin{aligned}\Pi &= -\frac{i\beta}{rb}\sum_{n=0}^{\infty}(2n+1)\{A_n\zeta_n(\xi_1)+[B_n+\zeta_n(b)]\psi_n(\xi_1)\}P_n(\cos\Theta)\\ &= -\frac{i\beta}{2rb}\sum_{0}^{\infty}(2n+1)\frac{\zeta_n(b)-k\eta_n(b)}{\zeta_n(x_1)-k\eta_n(x_1)}\{\zeta_n(x_1)\eta_n(\xi_1)-\eta_n(x_1)\zeta_n(\xi_1)\}P_n(\cos\Theta)\quad(\xi_1<b)\\ &= -\frac{i\beta}{2rb}\sum(2n+1)\Pi_n(\xi).P_n(\cos\Theta),\qquad\text{say}\end{aligned}\right\}\quad(41)$$

where

$$\Pi_n(\xi)=\frac{\zeta_n(b)-k\eta_n(b)}{\zeta_n(x_1)-k\eta_n(x_1)}\{\zeta_n(x_1)\eta_n(\xi_1)-\eta_n(x_1)\zeta_n(\xi_1)\}.\qquad(41\,\text{A})$$

4. *Approximate evaluation of the field*

To calculate the field at any point below the inversion it is necessary to transform the slowly convergent series (41) by complex integration (WATSONs transformation) of $s\,P_{s-\frac{1}{2}}(-\cos\Theta)\,\Pi_{s-\frac{1}{2}}(\xi)\sec s\pi$ quâ function of s around an infinite semi-circle bounded on the left by the imaginary s axis.

The layer functions $u_s(\xi)$, $v_s(\xi)$ {see (36) et seq} will be defined to reduce respectively to $\zeta_s(\xi)$ and $\eta_s(\xi)$ when $\varkappa\to 1$. Assuming that $\varkappa$ is an integral function of ξ, and writing (*c.f.* § 2.1)

$$\varkappa = 1+\delta_0\sum_0^{\infty}c_r\xi^r\qquad(42)$$

the equation satisfied by $u_{s-\frac{1}{2}}(\xi)$ and $v_{s-\frac{1}{2}}(\xi)$ is, by (31)

$$\frac{d^2\Omega}{d\xi^2}+\left\{1+\delta_0\sum_0^{\infty}c_r\xi^r-\frac{s^2-\frac{1}{4}}{\xi^2}\right\}\Omega=0.\qquad(43)$$

This equation has a regular singularity at $\xi=0$, the roots of the indicial equation being $\pm s+\frac{1}{2}$. It is then easy to show that the solution can be written

$$A\,\xi^{-s+\frac{1}{2}}\psi(-s,\xi)+B\,\xi^{s+\frac{1}{2}}\psi(s,\xi)\qquad(44)$$

where $\psi(-s,\xi)$* is an integral function of ξ. Since, when $\delta_0=0$, the solutions of (43) are $\xi^{\frac{1}{2}}J_{\pm s}(\xi)$, the function $\psi(-s,\xi)$ reduces to $\xi^s J_{-s}(\xi)$ when $\delta_0\to 0$.

Now

$$\zeta_{s-\frac{1}{2}}(\xi)=i\,\operatorname{cosec}s\pi\,.\,(\tfrac{1}{2}\pi\xi)^{\frac{1}{2}}\{J_{-s}(\xi)-e^{s\pi i}J_s(\xi)\}.\qquad(45)$$

Hence, by the definition of $u_s(\xi)$, we have,

$$u_{s-\frac{1}{2}}(\xi)=i\,\operatorname{cosec}s\pi\,(\tfrac{1}{2}\pi\xi)^{\frac{1}{2}}\{\xi^{-s}\psi(-s,\xi)-e^{s\pi i}\xi^s\psi(s,\xi)\}.\qquad(46)$$

This gives $u_{-s-\frac{1}{2}}(\xi)=e^{-s\pi i}.u_{s-\frac{1}{2}}(\xi)$, and similarly $v_{-s-\frac{1}{2}}(\xi)=e^{s\pi i}v^{s-\frac{1}{2}}(\xi)$. Thus $\Pi_{s-\frac{1}{2}}(\xi)$ is an even function of s, and since $P_{s-\frac{1}{2}}(-\cos\Theta)$ is also an even function of s, the function $s\,P_{s-\frac{1}{2}}(-\cos\Theta)\,\Pi_{s-\frac{1}{2}}(\xi).\sec s\pi$ is odd and so the integral along the imaginary axis vanishes. It is also easy to show, from (43), that as $|s|\to\infty$, ξ being fixed,

$$\left|\frac{\psi(s,\xi)}{J_s(\xi)}-1\right|=0\left(\frac{1}{s}\right).\qquad(47)$$

* There is no need to confuse this with the ψ of equation (33 A).

Hence

$$\left|\frac{u_{s-\frac{1}{3}}(\xi)}{\zeta_{s-\frac{1}{3}}(\xi)}-1\right|=0\left(\frac{1}{s}\right) \tag{47A}$$

and similarly, $v_{s-\frac{1}{3}}(\xi)\sim\eta_{s-\frac{1}{3}}(\xi)$, as $s\to\infty$. Therefore, $u_{s-\frac{1}{3}}(\xi)$ and $v_{s-\frac{1}{3}}(\xi)$ can be replaced by $\zeta_{s-\frac{1}{3}}(\xi)$ and $\eta_{s-\frac{1}{3}}(\xi)$ upon the infinite parts of the contour and it is thence possible to show that the contour integral converges.

Hence, by CAUCHY's theorem,

$$\left.\begin{aligned}\Pi=\frac{\pi i\beta}{rb}\sum\nu\;&\frac{\zeta_{\nu-\frac{1}{3}}(b)\,\eta_{\nu-\frac{1}{3}}(x_1)-\zeta_{\nu-\frac{1}{3}}(x_1)\,\eta_{\nu-\frac{1}{3}}(b)}{\eta_{\nu-\frac{1}{3}}(x_1)\,\frac{\partial}{\partial s}\{\zeta_{s-\frac{1}{3}}(x_1)-k_{s-\frac{1}{3}}\,\eta_{s-\frac{1}{3}}(x_1)\}_{(s=\nu)}}\times\\&\times\{\zeta_{\nu-\frac{1}{3}}(x_1)\,\eta_{\nu-\frac{1}{3}}(\xi_1)-\eta_{\nu-\frac{1}{3}}(x_1)\,\zeta_{\nu-\frac{1}{3}}(\xi_1)\}\times\sec\nu\pi\,.\,P_{\nu-\frac{1}{3}}(-\cos\Theta)\end{aligned}\right\}\tag{48}$$

the summation being over the roots of

$$\zeta_{s-\frac{1}{3}}(x_1)/\eta_{s-\frac{1}{3}}(x_1)=k \tag{49}$$

considered as an equation in s.

It can be shown (see Appendix II), that the field series (48) can be written

$$\left.\begin{aligned}\Pi=\frac{\pi i\beta}{rb}\sum\nu\;&\frac{\zeta_{\nu-\frac{1}{3}}(b)\,\eta_{\nu-\frac{1}{3}}(x_1)-\zeta_{\nu-\frac{1}{3}}(x_1)\,\eta_{\nu-\frac{1}{3}}(b)}{2i-\{\eta_{\nu-\frac{1}{3}}(x_1)\}^2\left(\frac{\partial k}{\partial s}\right)_{(s=\nu)}}\times\\&\times\{\zeta_{\nu-\frac{1}{3}}(x_1)\,\eta_{\nu-\frac{1}{3}}(\xi_1)-\eta_{\nu-\frac{1}{3}}(x_1)\,\zeta_{\nu-\frac{1}{3}}(\xi_1)\}\sec\nu\pi\times P_{\nu-\frac{1}{3}}(-\cos\Theta).\end{aligned}\right\}\tag{50}$$

It is worth noticing that the "reciprocity" theorem holds when a high level inversion is present, since Π is symmetrical in ξ and b.

If $b-x_1\ll x^{\frac{1}{3}}$; *i.e.* if $H<R^{\frac{1}{3}}\lambda^{\frac{2}{3}}$,—a condition satisfied during the Admiralty trials and for the lower terminal heights considered by SAXTON [9], (50) becomes*

$$\Pi=\frac{\lambda H z}{R^3\varkappa_0^{\frac{1}{2}}}\left(\frac{3\Gamma}{C}\right)^{\frac{1}{3}}\cdot\sum_{(\nu)}\frac{\nu P_{\nu-\frac{1}{3}}(-\cos\Theta)\sec\nu\pi}{1+\frac{1}{2}i\{\eta_{\nu-\frac{1}{3}}(x_1)\}^2\left(\frac{\partial k}{\partial s}\right)_{(s=\nu)}},\tag{51}$$

where z is the height of the receiving antenna.

When the inversion is absent (normal propagation), $k=0$, and the field series approximates to

$$\frac{\lambda H z}{R^3\varkappa_0^{\frac{1}{2}}}\left(\frac{3\Gamma}{C}\right)^{\frac{1}{3}}\sum\nu_0\,P_{\nu_0-\frac{1}{2}}(-\cos\Theta)\sec\nu_0\pi \tag{52}$$

where ν_0 stands for the roots of the equation in s, *viz.*,

$$\zeta_{s-\frac{1}{3}}(x_1)=0. \tag{53}$$

For all radio waves, x_1 is very large and the roots ν of (49) are such that $|\nu-x_1|$ is of order $x^{\frac{1}{3}}$; hence, arranging these roots in order of decreasing imaginary part *i.e.* writing

$$\nu_j=x_1+0(x^{\frac{1}{3}})-m_j\,x^{\frac{1}{3}}\,i \tag{54}$$

where $\{m_j\}$ is a certain positive increasing sequence, (it can be shown that $\Im_{(\nu)}$ is always negative), the dominant term of (51) sufficiently far from the transmitter

* Using the identity $\zeta\eta'-\zeta'\eta=2i$ and the fact that $r\doteq R$, $b\doteq 2\pi R\sqrt{\varkappa_0}/\lambda$, etc.

(the diffraction region) has a "Θ-factor" of $\Theta^{-\frac{1}{2}}\exp(-m_1\Theta)$. When the inversion is absent ($k=0$), the first term of series (52) is dominant at comparatively moderate distances ($> R^{\frac{1}{3}}\lambda^{\frac{1}{3}}$), but if one is present it turns out that the first few terms have to be considered even at the greatest ranges likely to be obtained in practice.

4.2. *The approximations to* $\zeta_s(\xi)$ and $\eta_s(\xi)$

The following approximations are fundamental in our discussion of the field series (51). When $|\xi-s|$ is not small compared with $x^{\frac{1}{3}}(\xi\sim x)$, they are given by (WATSON **9**, 1919; **12**, 1944)

$$\left.\begin{aligned} 2^{\frac{1}{2}}e^{-\frac{i\pi}{4}}\zeta_s(\xi) &\sim \{\xi/(\xi-s)\}^{\frac{1}{4}}\exp\{-2\sqrt{2}\,(\xi-s)^{\frac{3}{2}}i/3\,\xi^{\frac{1}{2}}\} \\ 2^{\frac{1}{2}}e^{\frac{i\pi}{4}}\eta_s(\xi) &\sim \{\xi/(\xi-s)\}^{\frac{1}{4}}\exp\{2\sqrt{2}\,(\xi-s)^{\frac{3}{2}}i/3\,\xi^{\frac{1}{2}}\} \end{aligned}\right\} \tag{55}$$

when $0<\arg\{(\xi-s)/\xi^{\frac{1}{2}}\}<\pi/3$, and by

$$\left.\begin{aligned} 2^{\frac{1}{2}}e^{-\frac{i\pi}{4}}\zeta_s(\xi) &\sim \{\xi/(\xi-s)\}^{\frac{1}{4}}[\exp\{-2\sqrt{2}\,(\xi-s)^{\frac{3}{2}}i/3\,\xi^{\frac{1}{2}}\}+i\exp\{2\sqrt{2}\,(\xi-s)^{\frac{3}{2}}i/3\,\xi^{\frac{1}{2}}\}] \\ 2^{\frac{1}{2}}e^{\frac{i\pi}{4}}\eta_s(\xi) &\sim \{\xi/(\xi-s)\}^{\frac{1}{4}}\exp\{2\sqrt{2}\,(\xi-s)^{\frac{3}{2}}i/3\,\xi^{\frac{1}{2}}\} \end{aligned}\right\} \tag{56}$$

when $\pi/3<\arg\{(\xi-s)/\xi^{\frac{1}{2}}\}<\pi$.

If, however, $|(\xi-s)/x^{\frac{1}{3}}|$ is small, the appropriate continuations of (55) and (56) are given by

$$\left.\begin{aligned} \zeta_s(\xi) &\sim \left\{\frac{\pi}{3}(\xi-s)\right\}^{\frac{1}{2}}e^{-\frac{i\pi}{6}}H_{\frac{1}{3}}^{(2)}\{2\sqrt{2}\,(\xi-s)^{\frac{3}{2}}/3\,\xi^{\frac{1}{2}}\} \\ \eta_s(\xi) &\sim \left\{\frac{\pi}{3}(\xi-s)\right\}^{\frac{1}{2}}e^{\frac{i\pi}{6}}H_{\frac{1}{3}}^{(1)}\{2\sqrt{2}\,(\xi-s)^{\frac{3}{2}}/3\,\xi^{\frac{1}{2}}\}. \end{aligned}\right\} \tag{57}$$

4.3. *Connection with ray theory*

As will appear later, the most important terms in series (51) are given by a few values of ν near x_1; *i.e.* writing

$$\nu = x(1+\varepsilon) \tag{58}$$

where $x=2\pi R/\lambda$

they are given by a few values of ε near* $\delta_0/2$. Writing also (see 39A),

$$y = x(1+\varrho+\tau/R), \qquad y_1 = x(1+\delta_0/2+\varrho) \tag{59}$$

where

$$\varrho = h/R \tag{59A}$$

and using the fact that $x\varrho^{\frac{3}{2}}\gg 1$ for all wavelengths and inversion heights likely to be of practical significance, the approximations (55), (56) give

$$\left.\begin{aligned} k = \frac{\left[N-L\{2(\varrho-\varepsilon)+\delta_0\}^{\frac{1}{2}}\left\{2(\varrho-\varepsilon)+\frac{2\tau}{R}\right\}^{\frac{1}{2}}\right]+i\left[M_2\left\{2(\varrho-\varepsilon)+\frac{2\tau}{R}\right\}^{\frac{1}{2}}+M_1\{2(\varrho-\varepsilon)+\delta_0\}^{\frac{1}{2}}\right]}{\left[N+L\{2(\varrho-\varepsilon)+\delta_0\}^{\frac{1}{2}}\left\{2(\varrho-\varepsilon)+\frac{2\tau}{R}\right\}^{\frac{1}{2}}\right]+i\left[M_2\left\{2(\varrho-\varepsilon)+\frac{2\tau}{R}\right\}^{\frac{1}{2}}-M_1\{2(\varrho-\varepsilon)+\delta_0\}^{\frac{1}{2}}\right]}\times \\ \times i\,\mathrm{Exp}\left[-\frac{4}{3}\sqrt{2}\,i\,x\left(p-\varepsilon+\frac{\delta_0}{2}\right)^{\frac{3}{2}}\right]. \end{aligned}\right\} \tag{60}$$

* This means $|\nu-x_1|\leq A\,x^{\frac{1}{3}}$; $\left|\varepsilon-\frac{\delta_0}{2}\right|\leq A\,x^{-\frac{1}{3}}$, where $A=0(1)$.

It will also appear later that $-\Im(\varepsilon)$ is always small compared with $\varrho - \Re(\varepsilon)$, for the dominant terms of (51); hence, by § 2,

$$k \doteqdot -i\,\overline{R} \exp\left\{-\frac{4}{3}\sqrt{2\,i\,x\left(\varrho - \varepsilon + \frac{\delta_0}{2}\right)^{\frac{3}{2}}}\right\} \tag{61}$$

where R is the complex reflection coefficient of a ray incident upon the underside of a layer of dielectric constant difference $\left(\delta_0 - \frac{2\tau}{R}\right)$ at an obliquity χ_1; and*

$$\chi_1^2 = 2(\varrho - \varepsilon_r) + \delta_0 \quad \text{where} \quad \varepsilon_r = \Re(\varepsilon). \tag{62}$$

Since $\varrho\, x^{\frac{2}{3}} \gg 1$, and $\left|\varepsilon - \frac{\delta_0}{2}\right|$ is $0(x^{-\frac{2}{3}})$, we have $\chi_1^2 \doteqdot 2\varrho$. (62A)

5. *"Thin" and "thick" layers*

The term "thin layer" will be applied to a layer sufficiently thin to behave like a dielectric discontinuity for a given wavelength. By §2, and (62A), it can be seen that the condition for the layer to behave approximately like a discontinuity is

$$\vartheta\sqrt{(2\varrho)} < 1 \quad \text{(approx.)}. \tag{63}$$

For values of τ and h liable to occur in practice this condition is often satisfied for metre waves but never for centimetre waves. When it is not satisfied and $\vartheta(2\varrho)^{\frac{1}{2}}$ becomes large it is apparent from § 2, and (62), that R becomes small. The inversion then has little effect upon propagation.

The most important case is therefore that when the layer is thin enough to behave like a discontinuity, and we investigate it below.

6. *Thin layers—metre wave propagation*

The equation for ν is (49) where k has its "discontinuity" value k_d, given by

$$k_d = \frac{\zeta'(y_1)\,\zeta(y) - \zeta'(y)\,\zeta(y_1)}{\eta'(y_1)\,\zeta(y) - \eta(y_1)\,\zeta'(y)}. \tag{64}$$

In approximating to the field series (51) it is important to note that each term is made up of two factors, namely,

(A) $\left[1 + \frac{1}{2}\,i\,\{\eta_{\nu_j}(x_1)\}^2 \left(\frac{\partial k}{\partial s}\right)_{(\nu_j)}\right]^{-1} = F_j^{-1}$, say, and

(B) A factor $\Theta^{-\frac{1}{2}} \exp(-m_j\Theta)$, arising from the factor $\nu_j P_{\nu_j - \frac{1}{2}}(-\cos\Theta)\sec\nu_j\pi$.

It is shown in the Appendix that we have the following approximations for the factor F.

$$F \doteqdot 1 + \left(\delta_0 - \frac{2\tau}{R}\right)\{\eta_{\nu-\frac{1}{2}}(x_1)\}^2\{\psi_{\nu-\frac{1}{2}}(y_1)\}^2 \tag{65}$$

if

$$x^{\frac{2}{3}}\left|\varepsilon - \frac{\delta_0}{2}\right| \gg 1, \qquad \Re\left\{x^{\frac{2}{3}}\left(\varepsilon - \frac{\delta_0}{2}\right)\right\} \gg 1$$

and

$$F \doteqdot 1 - \sqrt{2\left(\varrho - \varepsilon + \frac{\delta_0}{2}\right)^{\frac{1}{2}}}\,\eta_{\nu-\frac{1}{2}}(x_1)\,\zeta_{\nu-\frac{1}{2}}(x_1) \tag{66}$$

provided that ε is not too near ϱ or $\varrho - \frac{\delta_0}{2}$.

* See also Appendix IV.

VOL. 2 1952

Using approximations (55 etc.), it follows that, if $x^{\frac{2}{3}}\left(\varepsilon-\frac{\delta_0}{2}\right)$ is not small and $\Re\left\{x^{\frac{2}{3}}\left(\varepsilon-\frac{\delta_0}{2}\right)\right\}>1$

$$F \doteqdot 1-i\left\{\left(\varrho-\varepsilon+\frac{\delta_0}{2}\right)\Big/\left(\frac{\delta_0}{2}-\varepsilon\right)\right\}^{\frac{1}{2}}\left[\exp\left\{\frac{4}{3}\sqrt{2}\,x\,i\left(-\varepsilon+\frac{\delta_0}{2}\right)^{\frac{3}{2}}\right\}-1\right] \tag{67}$$

and this is large of order

$$\left\{\left(\varrho-\varepsilon+\frac{\delta_0}{2}\right)\Big/\left(\frac{\delta_0}{2}-\varepsilon\right)\right\}^{\frac{1}{2}}\exp\left\{\frac{4}{3}\sqrt{2}\,x\,i\left(\varepsilon-\frac{\delta_0}{2}\right)^{\frac{3}{2}}\right\}.$$

If $\Re\left\{x^{\frac{2}{3}}\left(\varepsilon-\frac{\delta_0}{2}\right)\right\}<-1$, F is approximately equal to,

$$1-\left\{\left(\varrho-\varepsilon+\frac{\delta_0}{2}\right)\Big/\left(-\varepsilon+\frac{\delta_0}{2}\right)\right\}^{\frac{1}{2}}. \tag{68}$$

Finally, if $x^{\frac{2}{3}}\left(\varepsilon-\frac{\delta_0}{2}\right)$ is small, F is approximately,

$$1-1{\cdot}780\,x^{\frac{1}{3}}\left(\varrho-\varepsilon+\frac{\delta_0}{2}\right)^{\frac{1}{2}}\left\{1+0{\cdot}9181\,x^{\frac{2}{3}}\left(\varepsilon-\frac{\delta_0}{2}\right)+0{\cdot}8429\,x^{\frac{4}{3}}\left(\varepsilon-\frac{\delta_0}{2}\right)^2+\cdots\right\}. \tag{69}$$

Expression (69) gives a good approximation to F when $x^{\frac{2}{3}}\left|\left(\varepsilon-\frac{\delta_0}{2}\right)\right|<1$.

Having regard to the characteristic values ν_j presently to be determined, it can now be shown that only those for which $x^{\frac{2}{3}}\,\varepsilon$ is not large give significant contributions to the field series (51).

7. *The characteristic roots*

7.1. Using the approximations (55–57), and writing $s=x(1+\varepsilon)$, $x=2\pi R/\lambda$, $\delta_0=\varkappa_0-1$, {*c.f.* (58), (35A) and (7)} and putting $k=k_d$, we find that equation (49) takes the following forms.

(1) $$\operatorname{Exp}\left\{-\frac{4}{3}\sqrt{2}\,i\,x\left(-\varepsilon+\frac{\delta_0}{2}\right)^{\frac{3}{2}}\right\}=i\,(1+k_a) \tag{70}$$

when $\frac{4\pi}{3}<\arg\left(\varepsilon-\frac{\delta_0}{2}\right)<2\pi$

(2) $$i\operatorname{Exp}\left\{-\frac{4}{3}\sqrt{2}\,i\,x\left(-\varepsilon+\frac{\delta_0}{2}\right)^{\frac{3}{2}}\right\}=k_d \tag{71}$$

when $\pi<\arg\left(\varepsilon-\frac{\delta_0}{2}\right)<\frac{4\pi}{3}$, (71A)

and

$$e^{-\frac{i\pi}{3}}\,\frac{H_{\frac{1}{3}}^{(2)}\left\{\frac{2\sqrt{2}}{3}\,x\left(-\varepsilon+\frac{\delta_0}{2}\right)^{\frac{3}{2}}\right\}}{H_{\frac{1}{3}}^{(1)}\left\{\frac{2\sqrt{2}}{3}\,x\left(-\varepsilon+\frac{\delta_0}{2}\right)^{\frac{3}{2}}\right\}}=k_d \tag{72}$$

when $x\left|-\varepsilon+\frac{\delta_0}{2}\right|^{\frac{3}{2}}$ is small.

If $k_d=0$, the inversion disappears, (normal propagation) and we then have the value ε_0 of ε corresponding to the dominant term of (52) in the diffraction region, *viz.*,

$$\varepsilon_0=\frac{\delta_0}{2}+1{\cdot}855\,x^{-\frac{2}{3}}e^{\frac{5i\pi}{3}}. \tag{73}$$

7.2. *The roots of* (70). $(X > 1)$

Let

$$x^{\frac{2}{3}}\left(\varepsilon - \frac{\delta_0}{2}\right) = X - iY \tag{74}$$

then

$$x^{\frac{2}{3}}\left(\varrho - \varepsilon + \frac{\delta_0}{2}\right) = P - X + iY, \qquad x^{\frac{2}{3}}(\varrho - \varepsilon) = P - X - Z + iY \tag{75}$$

where

$$P = x^{\frac{2}{3}}\varrho, \qquad Z = x^{\frac{2}{3}}\left(\frac{\delta_0}{2}\right). \tag{76}$$

Since $P \geq 1$, we can obtain a solution for which $Y \ll X$, $Y \ll P - X - Z$, $P > X$. From expressions (55) we have

$$\left.\begin{aligned} \zeta'(y_1)/\zeta(y_1) &\sim -i\{2(\varrho - \varepsilon) + \delta_0\}^{\frac{1}{2}} \\ &\sim -i\,x^{-\frac{1}{3}}\{2(P - X + iY)\}^{\frac{1}{2}} \end{aligned}\right\} \tag{77}$$

$$\eta'(y_1)/\eta(y_1) \sim i\,x^{-\frac{1}{3}}\{2(P - X + iY)\}^{\frac{1}{2}} \tag{78}$$

$$\left.\begin{aligned} \zeta'(y)/\zeta(y) &\sim -i\{2(\varrho - \varepsilon)\}^{\frac{1}{2}} \\ &\sim -i\,x^{-\frac{1}{3}}\{2(P - X - Z + iY)\}^{\frac{1}{2}}. \end{aligned}\right\} \tag{79}$$

Also

$$S(y_1)/\eta(y_1) \sim i\exp\left\{-\frac{4}{3}\sqrt{2}\,i\,x\left(\varrho - \varepsilon + \frac{\delta_0}{2}\right)^{\frac{3}{2}}\right\} \tag{80}$$

$$\sim i\exp\{-\tfrac{4}{3}\sqrt{2}\,i(P - X + iY)^{\frac{3}{2}}\} \tag{81}$$

and

$$\exp\left\{-\frac{4}{3}\sqrt{2}\,x\left(-\varepsilon + \frac{\delta_0}{2}\right)^{\frac{3}{2}}i\right\} = \exp\left\{-\frac{4}{3}\sqrt{2}\,(X - iY)^{\frac{3}{2}}\right\} \ll 1 \tag{82}$$

if X is much greater than unity. Assuming $X > 1$, and using equations (77–82), equation (70) becomes, approximately

$$-1 = -\frac{iZ}{4(P - X - Z + iY)}\exp\left[-\frac{4}{3}\sqrt{2}\,i\left\{(P - X)^{\frac{3}{2}} + \frac{3}{2}iY(P - X)^{\frac{1}{2}}\right\}\right] \tag{83}$$

provided that

$$Z/(P - X - Z) < \tfrac{1}{3}, \qquad \text{say} \tag{83A}$$

Hence, approximately,

$$-\frac{4}{3}\sqrt{2}\,i(P - X)^{\frac{3}{2}} + 2\{2(P - X)\}^{\frac{1}{2}}Y = -\left(2p + \frac{1}{2}\right)i\pi + \log\left\{\frac{4(P - X - Z)}{Z} + \frac{4iY}{Z}\right. \tag{84}$$

where p is an integer.

Equating imaginary parts,

$$P - X = \tfrac{1}{2}\sigma_p, \qquad \text{approximately} \tag{85}$$

where

$$\sigma_p^3 = 3\pi(p + \tfrac{1}{4}). \tag{85A}$$

Hence by (83A)

$$p + \tfrac{1}{4} > 2{\cdot}4\,Z^{\frac{3}{2}}. \tag{86}$$

Equating real parts and using (85), we find

$$Y_p = \frac{1}{2\sqrt{\sigma_p}}\log\left(\frac{2\sigma_p}{Z} - 4\right) \tag{87}$$

VOL. 2 1952

provided that $Y \ll X$, (say $< X/3$); *i.e.*, provided that

$$\frac{3}{2\sqrt{\sigma_p}} \log\left(\frac{2\sigma_p}{Z} - 4\right) < P - \frac{1}{2}\sigma_p . \tag{88}$$

It will appear below that p has to be taken as large as is compatible with a certain condition. The left of (88) is then small compared with unity for all values of Z likely to occur in practice. Thus, to satisfy (88) and the condition $X > 1$, we must have

$$\sigma_p < 2(P-1) . \tag{89}$$

The permissible values of p are therefore given by the inequality,

$$2{\cdot}4\, Z^{\frac{2}{3}} < p + \tfrac{1}{4} < 0{\cdot}3(P-1)^{\frac{2}{3}} . \tag{90}$$

7.3. *The roots of* (71). $\{X < -1\}$

Since Y is always positive and (as will appear later) $\ll X$, the condition (71 A) implies that X is negative. We can conveniently suppose that $X < -1$. (The case $-1 < X < 1$ will be considered presently.)

Using approximations (55–57) equation (71) becomes

$$\left.\begin{aligned} &\exp\left\{-\tfrac{4}{3}\sqrt{2}\, i(-X)^{\frac{3}{2}} + 2Y(-2X)^{\frac{1}{2}}\right\} \\ &\qquad = -\frac{Z}{4(P-X-Z+iY)} \exp\left[-\tfrac{4}{3}\sqrt{2}\, i(P-X)^{\frac{3}{2}} + 2Y\{2(P-X)\}^{\frac{1}{2}}\right] \end{aligned}\right\} \tag{91}$$

approximately.

Therefore,

$$\tfrac{4}{3}\sqrt{2}\{(P-X)^{\frac{3}{2}} - (-X)^{\frac{3}{2}}\} = (2r+1)\pi \tag{92}$$

where, since $X < -1$, r is any integer greater than the greatest integer for which

$$\tfrac{4}{3}\sqrt{2}\{(P+1)^{\frac{3}{2}} - 1\} > (2r+1)\pi . \tag{93}$$

Let $X_r (r > r_0)$ be the values of X defined by (92). Then approximately,

$$Y_r = \frac{\log\left\{\dfrac{4(P-X_r)}{Z} - 4\right\}}{2\left[\{2(P-X_r)\}^{\frac{1}{2}} - \{-2X_r\}^{\frac{1}{2}}\right]} . \tag{94}$$

Hence, when r is made so large that $-X_r \gg P$, Y_r becomes large and therefore only a finite number of the terms contributed by the roots of (71) to the residue series (51) are significant. Usually, however, only the first few of these are important at distances exceeding the critical distance.

7.4. *The roots of* (72).

If $|X + iY| < 1$, say, we find, from (57), the approximate equation,

$$\frac{\zeta_s(x_1)}{\eta_s(x_1)} = e^{\frac{2i\pi}{3}} \frac{1 + e^{\frac{i\pi}{3}} 0{\cdot}9181(X - iY) + \frac{1}{3}(X-iY)^3 + 0\{(X-iY)^4\}}{1 + e^{-\frac{i\pi}{3}} 0{\cdot}9181(X - iY) + \frac{1}{3}(X-iY)^3 + 0\{(X-iY)^4\}} \tag{95}$$

so that,

$$\log \frac{\zeta_s(x_1)}{\eta_s(x_1)} \doteq \frac{2i\pi}{3} + 1{\cdot}59\,i(X - i\,Y) - 0{\cdot}7300\,i(X - i\,Y)^2 + 0\{(X - i\,Y)^4\}. \quad (96)$$

Hence, approximately,

$$\left.\begin{aligned} &\exp\left[\frac{2i\pi}{3} + 1{\cdot}59(i\,X + Y) - 0{\cdot}73\{2\,X\,Y + i\,(X^2 - Y^2)\}\right] \\ &\quad = -\frac{i\,Z}{4(P - X - Z - i\,Y)} \exp\left[-\frac{4}{3}\sqrt{2}\,i\,(P - X)^{\frac{3}{2}} + 2\,Y\,\{2(P - X)\}^{\frac{1}{2}}\right]. \end{aligned}\right\} \quad (97)$$

Equating real and imaginary parts we have the approximate equations,

$$\tfrac{4}{3}\sqrt{2}\,(P - X)^{\frac{3}{2}} + 1{\cdot}59\,X - 0{\cdot}73(X^2 - Y^2) = (2K + \tfrac{5}{8})\,\pi \quad (98)$$

and

$$1{\cdot}59\,Y - 1{\cdot}46\,X\,Y = -\log\left\{\frac{4(P - X)}{Z} - 4\right\} + 2\,Y\,\{2(P - X)\}^{\frac{1}{2}} \quad (98\,\mathrm{A})$$

giving,

$$Y = \frac{\log\left\{\dfrac{4(P - X)}{Z} - 4\right\}}{2\,\{2(P - X)\}^{\frac{1}{2}} - 1{\cdot}59 + 1{\cdot}46\,X} \quad (99)$$

The term $0{\cdot}73\ (X^2 - Y^2)$ in (98) will normally be small compared with the other terms. We proceed tentatively, obtaining (for a given values of P), a first approximation to X, X', say, by solving the equation

$$\tfrac{4}{3}\sqrt{2}\,(P - X)^{\frac{3}{2}} + 1{\cdot}59\,X = (2K + \tfrac{5}{8})\,\pi \quad (100)$$

where K is an integer (or integers) chosen to make $|X| < 1$. Denoting by Y' the corresponding value of Y, which is then given by (99), the second approximation X'' is obtained from

$$\tfrac{4}{3}\sqrt{2}\,(P - X)^{\frac{3}{2}} + 1{\cdot}59\,X - 0{\cdot}73\,(X'^2 - Y'^2) = (2K + \tfrac{5}{8})\,\pi \quad (101)$$

and so on. In practice, the calculations are somewhat simplified by the fact that P is large.

8. *Series for the field*

Using LAPLACEs approximation for $P_{\nu-\frac{1}{2}}(-\cos\Theta)$ (see WATSONs paper) it now appears that the series (51) is equivalent to the series,

$$\frac{2Hz}{R^2}\left(\frac{3\,\Gamma\lambda}{C\,d}\right)^{\frac{1}{2}} \sum_1^\infty \frac{1}{F_j} \exp\{-(X_j\,i + Y_j)\,x^{\frac{1}{2}}\,\Theta\} \quad (102)$$

where $d = R\Theta$, and X_j and Y_j are defined in terms of ν_j and ε_j by (74); their values being so arranged that $\{Y_j\}$ is an increasing positive sequence. Hence, by (26 A) the series for the electric field E_ϕ is

$$\frac{8\pi^2\sqrt{\varkappa_0}\,Hz}{R\,\lambda^{\frac{3}{2}}}\left(\frac{3\,\Gamma}{C\,d}\right)^{\frac{1}{2}} \sum_1^\infty \frac{1}{F_j} \exp\{-(X_j\,i + Y_j)\,x^{\frac{1}{2}}\,\Theta\}. \quad (103)$$

VOL. 2 1952

The values of F_j are, by (66–69), as follows:

$$F_j \doteqdot 1 - 1{\cdot}780\,(P - X_j + i\,Y_j)^{\frac{1}{2}}\{1 + 0{\cdot}9181\,(X_j - i\,Y_j) + 0{\cdot}8429\,(X_j - i\,Y_j)^2\}, \quad (104)$$

if $|X_j + i\,Y_j| < 1$,

$$F_j \doteqdot 1 - \{(P - X_j + i\,Y_j)/(-X_j + i\,Y_j)\}^{\frac{1}{2}}, \quad (105)$$

if $X_j < -1$, say,

$$F_j \doteqdot 1 + \{(P - X_j + i\,Y_j)/(X_j - i\,Y_j)\}^{\frac{1}{2}}\,[\exp\{\tfrac{4}{3}\sqrt{2}\,(X_j - i\,Y_j)^{\frac{3}{2}}\} - i], \quad (106)$$

if $X_j > 1$, say.

This series has to be compared with the corresponding series for the case of normal propagation; the latter is, however, so highly convergent at distances exceeding $R^{\frac{2}{3}}\lambda^{\frac{1}{3}}$, that its first term then gives an excellent approximation to its sum. By (52), (73) and (26 A), the approximate value of (103) when the propagation is normal*, is, when $d > R^{\frac{2}{3}}\lambda^{\frac{1}{3}}$.

$$\frac{8\pi^2\sqrt{\varkappa_0}\,.\,Hz}{R\lambda^{\frac{3}{2}}}\left(\frac{3\,\Gamma}{C\,d}\right)^{\frac{1}{2}}\mathrm{Exp}\{-(0{\cdot}9272\,i + 1{\cdot}606)\,\Theta\,x^{\frac{1}{2}}\}. \quad (107)$$

In subsequent numerical work we shall sometimes require the normal field strength at distances somewhat less than $R^{\frac{2}{3}}\lambda^{\frac{1}{3}}$; we therefore give the second term of the normal field series—it can be shown to be

$$\frac{8\pi^2\sqrt{\varkappa_0}\,.\,Hz}{R\lambda^{\frac{3}{2}}}\left(\frac{3\,\Gamma}{C\,d}\right)^{\frac{1}{2}}\mathrm{Exp}\{-(1{\cdot}622\,i + 2{\cdot}809)\,\Theta\,x^{\frac{1}{2}}\}. \quad (108)$$

We now illustrate the above theory by applying it to the CARDINGTON inversion (see § 2.2). For this inversion, $(\Delta\mu)_0 = 2{\cdot}56 \times 10^{-5}$, and $h = 1200$ m. Following SAXTON, we allow for the normal slow lapse rate of μ with height in the atmosphere below and above the inversion (assumed standard,) by taking for the earth's radius R, an "effective" value of 4/3 times the actual value. The most important values of X_j, Y_j, and $1/F_j$, for the cases $\lambda = 3\frac{1}{2}$ m and $\lambda = 7$ m are found to be

$\lambda = 3\frac{1}{2}$ m

X	1·106,	0·245,	−0·726,	−1·86,	−3·11,	−4·524,	−6·084,
Y	0·348,	0·408,	0·495,	0·585,	0·687,	0·792,	0·900,
$-\lvert F_j^{-1}\rvert$	−0·04167,	0·2024,	0·4378,	0·7463,	1·072	1·43,	1·808,
$\arg(I/F_j)$	50° 46′,	3° 6′,	−23° 23′	12° 27′,	9° 26′,	7° 54′,	6° 56′,

X	7·724,	9·554.
Y	1·007,	1·12.
$-\lvert F_j^{-1}\rvert$	2·20,	2·56.
$\arg(I/F_j)$	6° 14′,	5° 2′.

* We have omitted from (103) and (107), a common phase factor $\mathrm{Exp}\{-i\,x_1\,\Theta\}$.

$\lambda = 7\ m$*

X	1·636,	0·518,	— 0·724,	— 2·4,	— 4·43.
Y	0·438,	0·505,	0·691,	0·932,	1·032.
$-\lvert F_j^{-1}\rvert$	— 0·0139,	0·165,	0·619,	1·294,	2·063.
$\arg(I/F_j)$	78° 28′,	32° 47′,	— 51° 24′,	16° 37′,	10° 59′.

The curvilinear attenuation $Y_j x^{\frac{1}{2}}$ increases steadily as $-X_j$ increases, but so also does $|F_j^{-1}|$. The effect of this is to make the series (103) sufficiently convergent for the sum of the terms given by the above values of X_j and Y_j to give a good approximation to its sum, only at distances exceeding about 160 km.

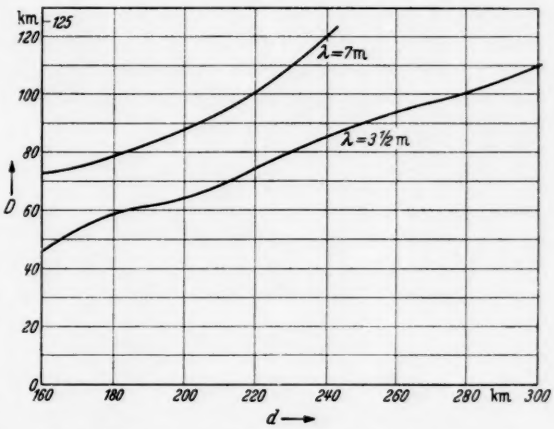


Fig. 1. Equivalent distance graph for CARDINGTON Inversion. $(\Delta\mu)_0 = 2{\cdot}58 \times 10^{-5}$, $h = 1200$ m.

Equivalent distance

The magnitudes of (103) and (107) may conveniently be compared by means of the concept of "equivalent distance." We define the "equivalent distance" D, corresponding to an actual transmitter-receiver distance d, as the distance at which the *normal* field strength would be equal to the actual field strength received at distance d with the inversion present. It is clear from the figure that an inversion such as that measured at CARDINGTON in 1945 can cause considerable anomalous propagation upon the above wavelengths.

9. *Note on the applicability of the theory*

The present theory assumes that the inversion layer is a complete spherical surface concentric with the earth's surface. In practice, however, it is of limited extent only. Unfortunately, any attempts to modify the analysis accordingly lead to very formidable mathematical difficulties whose elucidation seems to be beyond the scope of present technique. It is, nevertheless, possible to gain some idea of the conditions under which the theory can be expected to be approximately true. At distances near to, or beyond, the limiting distance, the first few modes in the field series (103) are dominant: by (62), these modes correspond to rays of obliquity

$$\{2(\varrho - \varepsilon_r) + \delta_0\}^{\frac{1}{2}} \doteqdot (2\varrho)^{\frac{1}{2}} - \frac{\varepsilon_r}{2(2\varrho)^{\frac{1}{2}}}$$

where $\varepsilon_r = \Re(\varepsilon) = 0(x^{-\frac{3}{2}})$ for these principal modes.

* The case $\lambda = 7$ m has an interesting application to television.

VOL. 2 1952

These rays form a pencil of angular crossection

$$\frac{1}{2(2\varrho)^{\frac{1}{2}}} \cdot \text{Max}\,|\varepsilon_\tau| = \frac{A\,x^{-\frac{3}{2}}}{2(2\varrho)^{\frac{1}{2}}}$$

where A is a pure number of order unity. By easy geometry, these rays strike the underside of the inversion in a region above the semi-limiting distance whose linear dimensions are of order $Rx^{-\frac{3}{2}}/2(2\varrho)^{\frac{1}{2}}$; *i.e.* of order $0{\cdot}1R^{\frac{2}{3}}\lambda^{\frac{1}{3}}/h^{\frac{1}{2}}$ $(= W$ say).

It seems, therefore, that the main contribution to the field near to, or beyond, the limiting distance, is due to diffraction from rays reflected from this region. Hence the present theory can probably be applied to discuss the field at such distances, provided that the inversion amply covers this critical region. For the CARDINGTON inversion, $W = 1{\cdot}72\,\lambda^{\frac{2}{3}}$ km (λ in metres), but even for an inversion as high as 5000 m (very rare), $W = 0{\cdot}84\,\lambda^{\frac{2}{3}}$ km. Hence, from what is known of high level inversions (at any rate around the British Isles) it is unlikely that they are ever of sufficient extent to affect propagation on wavelengths much longer than about 10 m.

(Note: Even if the inversion were of sufficient extent for the theory to be applicable to wavelengths of many metres, there would be little anomalous propagation, since it can be shewn that, if $P < 0{,}8$, propagation approximates closely to normal.)

10. *Criteria for "good" anomalous propagation near to, and beyond, the limiting distance.*

By § 9, and (63), these are:

(i) The inversion must be well developed in a region above the mid-point of the limiting distance whose dimensions are not small compared with $0{\cdot}1\,R^{\frac{2}{3}}\lambda^{\frac{1}{3}}/h^{\frac{1}{2}}$.

(ii) The inversion must be thin enough to behave like a discontinuity, *i.e.* we must have (see 63),

$$\tau < 329\,\lambda/\sqrt{h} \qquad (109)$$

where λ and h are in metres.

It is easy to see that (ii) is satisfied (for metre waves) by the CARDINGTON inversion* and, as far as is known, this inversion was of sufficient extent to satisfy (1).

11. CONCLUSION

The theory which has been developed shows that metre wave anomalous propagation can be caused by high level subsidence inversions but the theory needs further experimental verification. In the writer's opinion, this might usefully take the form of attempted correlation between field strength and inversion measurements. The latter should, where possible, be made with the special apparatus already referred to in § 2·2, since Radio-Sonde technique is insufficiently sensitive to measure the fine structure of those inversions which are sufficiently sharp to cause significant anomalous propagation.

On the theoretical side, it might be desirable to compute "Equivalent Distance" graphs for other high level inversions, and to investigate the field strength at large heights with a view to possible applications to aircraft.

* For this inversion (109) is $\tau < 9{\cdot}5\,\lambda$ metres.

Acknowledgements—The writer is indebted to L. E. BEGHIAN, at whose suggestion this investigation was undertaken, for his kind assistance during the early stages of the work, to the Captain-Superintendent, A.S.R.E., for permission to refer to the Admiralty trials, to A.T. PRICE for many helpful discussions, and to his wife for checking the calculations.

APPENDIX I

The radiation field of a horizontal electric dipole

Defining Λ_0 as in the text and introducing the vector

$$\Lambda_1 = (R+H)\{\varkappa_0 \Pi_0 \cos\Theta, -\varkappa_0 \Pi_0 \sin\Theta, 0\},$$

which is in the direction of the vertical at the dipole position, then

$$\Lambda_1 - \Lambda_0 = \frac{\beta C}{ip}\sqrt{\varkappa_0}\,\mathrm{grad}\left\{\mathrm{Exp}\left(-\frac{ip}{C}\sqrt{\varkappa_0}\,R_1\right)\right\} \tag{a}$$

so that

$$\mathrm{curl}\,\Lambda_1 = \mathrm{curl}\,\Lambda_0. \tag{b}$$

Hence the undisturbed field is equivalent to that produced by a vertical magnetic dipole; transposing the origin to the dipole position for convenience, using Λ_1 instead of Λ_0 in the equations (26), and remembering that Π_0 is now $\frac{\beta}{r}\exp\left(-\frac{ip}{C}\sqrt{\varkappa_0}\,r\right)$, we find that the radiation field of this dipole is,

$$H_r \simeq 0\left(\frac{1}{r^2}\right); \qquad H_\Theta \sim -(R+H)\frac{p^2}{C^2}\varkappa_0 \sin\Theta\,.\,\Pi_0; \qquad H_\Phi = 0$$

$$E_r = 0; \qquad E_\Theta = 0; \qquad E_\Phi \sim (R+H)\frac{p^2}{C^2}\sqrt{\varkappa_0}\sin\Theta\,.\,\Pi_0. \tag{c}$$

For a horizontal electric dipole with its axis in the direction $\Phi = \frac{\pi}{2}$, the Hertzian vector is

$$\Lambda_2 = (R+H)\{\sqrt{\varkappa_0}\,\Pi_0 \sin\varphi\sin\Theta, \sqrt{\varkappa_0}\,\Pi_0 \sin\varphi\cos\Theta, \sqrt{\varkappa_0}\,\Pi_0\cos\varphi\}$$

where

$$\Pi_0 = \frac{\beta}{r}\,\mathrm{Exp}\left(-\frac{ip\sqrt{\varkappa_0}}{C}\,r\right)$$

and the field is now given by

$$\mathsf{E} = \frac{1}{\varkappa_0}\,\mathrm{curl}\,\mathrm{curl}\,\Lambda_2, \qquad \mathsf{H} = \frac{ip}{C}\,\mathrm{curl}\,\Lambda_2$$

and the "radiation" field by

$$H_r \sim 0\left(\frac{1}{r^2}\right); \quad H_\Theta = -(R+H)\frac{p^2}{C^2}\varkappa_0\cos\varphi\,.\,\Pi_0; \quad H_\Phi \sim (R+H)\frac{p^2}{C^2}\varkappa_0\sin\varphi\cos\Theta\,.\,\Pi_0$$

$$E_r \sim 0\left(\frac{1}{r^2}\right); \quad E_\Theta = (R+H)\frac{p^2}{C^2}\sqrt{\varkappa_0}\sin\varphi\cos\Theta\,.\,\Pi_0; \quad E_\Phi \sim (R+H)\frac{p^2}{C^2}\sqrt{\varkappa_0}\cos\varphi\,.\,\Pi_0. \tag{d}$$

Putting $\varphi = 0$, this agrees with the field produced by the vertical magnetic dipole if $\Theta \doteq \frac{\pi}{2}$; *i.e.* under the conditions stated in the text.

Let Γ be the radiated power in theoretical units; then Γ equals the mean value* of

$$\frac{c}{4\pi}\iint \mathsf{E}_{\Delta}\mathsf{H}\,.\,d\mathsf{S}$$

taken over a large closed surface surrounding the dipole. Evaluating this integral either (c) or (d) give

$$\Gamma \doteqdot \frac{C}{3}\beta^2 R^2\left(\frac{2\pi}{\lambda}\right)^4 \varkappa_0^{\frac{3}{2}} \qquad (R + H \doteqdot R).$$

Hence

$$\beta = \frac{\lambda^2}{4\pi^2 R}\sqrt{\left(\frac{3\Gamma}{C\,\varkappa_0^{\frac{3}{2}}}\right)}.$$

APPENDIX II(A)

The factor $\eta_{\nu-\frac{1}{2}}(\varkappa_1)\frac{\partial}{\partial s}\{\zeta_{s-\frac{1}{2}}(\varkappa_1) - k\,\eta_{s-\frac{1}{2}}(\varkappa_1)\}_{s=\nu}$.

Since $\xi - x$ is $0(x^{\frac{1}{3}})$,

$$2\sqrt{2}\,(\xi - s)^{\frac{3}{2}}/3\,\xi^{\frac{1}{2}} = \{2\sqrt{2}\,(\xi - s)^{\frac{3}{2}}/3\,x^{\frac{3}{2}}\}\{1 + 0\,(x^{-\frac{3}{2}})\}.$$

Hence, since x is always extremely large, we may write, with negligible error,

$$2\sqrt{2}\,(\xi - s)^{\frac{3}{2}}/3\,\xi^{\frac{1}{2}} = 2\sqrt{2}\,(\xi - s)^{\frac{3}{2}}/3\,x^{\frac{1}{2}}$$

in approximations (57). Hence, approximately,

$$\frac{\partial}{\partial s}\{\zeta_s(\xi)\} = -\,\zeta_s'(\xi); \qquad \frac{\partial}{\partial s}\{\eta_s(\xi)\} = -\,\eta_s'(\xi)$$

the dash denoting $\partial/\partial\xi$. We have, therefore,

$$\eta_{\nu-\frac{1}{2}}(\xi)\frac{\partial}{\partial s}\{\zeta_{s-\frac{1}{2}}(\xi) - k\,\eta_{s-\frac{1}{2}}(\xi)\}_{(\nu)} = \eta_{\nu-\frac{1}{2}}(\xi)\left[\frac{\partial}{\partial s}\{\zeta_{s-\frac{1}{2}}(\xi)\} - k\,\frac{\partial}{\partial s}\{\eta_{s-\frac{1}{2}}(\xi)\}\right]_{(s=\nu)} -$$

$$-\{\eta_{\nu-\frac{1}{2}}(\xi)\}^2\left(\frac{\partial k}{\partial s}\right)_{(s=\nu)},$$

$$\doteqdot -\,\eta_{\nu-\frac{1}{2}}(\xi)\,[\zeta'_{\nu-\frac{1}{2}}(\xi) - k\,\eta'_{\nu-\frac{1}{2}}(\xi)] - \{\eta_{\nu-\frac{1}{2}}(\xi)\}^2\left(\frac{\partial k}{\partial s}\right)_{s=\nu}$$

Using the relations,

$$\zeta_{\nu-\frac{1}{2}}(x_1)/\eta_{\nu-\frac{1}{2}}(x_1) = k, \qquad \zeta_s(x_1)\,\eta_s'(x_1) - \zeta_s'(x_1)\,\eta_s(x_1) = 2\,i$$

we have,

$$\eta_{\nu-\frac{1}{2}}(x_1)\frac{\partial}{\partial s}\{\zeta_{s-\frac{1}{2}}(x_1) - k\,\eta_{s-\frac{1}{2}}(x_1)\}_{s=\nu} \doteqdot 2\,i - \{\eta_{\nu-\frac{1}{2}}(x_1)\}^2\left(\frac{\partial k}{\partial s}\right)_{(\nu)} = 2\,i\,F_j.$$

APPENDIX III(B)

The factor F_j

We now require an approximation to this factor when

$$k = k_d = \frac{\zeta'(y_1)\,\zeta(y) - \zeta'(y)\,\zeta(y_1)}{\eta'(y_1)\,\zeta(y) - \eta(y_1)\,\zeta'(y)}.$$

---

* Π_0 has here to be taken as $\frac{\beta}{r}\cos\left\{\frac{p}{C}\sqrt{\varkappa_v}\,(r - ct)\right\}$.

We have*,

$$\frac{\partial k}{\partial s} \doteq -\frac{\partial k}{\partial \xi}$$

$$= -\frac{\zeta''(y_1)\,\zeta(y) - \zeta''(y)\,\zeta(y_1)}{\zeta(y)\,\eta'(y_1) - \zeta'(y)\,\eta(y_1)} + \frac{\{\zeta'(y_1)\,\zeta(y) - \zeta'(y)\,\zeta(y_1)\}\{\zeta(y)\,\eta''\zeta(y_1) - \zeta''(y)\,\eta(y_1)\}}{\{\zeta(y)\,\eta'(y_1) - \zeta'(y)\,\eta(y_1)\}^2}$$

Using the relations $\zeta''(\xi) + \left(1 - \frac{s^2}{\xi^2}\right)\zeta(\xi) = 0$, etc., and the identity $\zeta\eta' - \zeta'\eta = 2i$, this reduces to

$$\frac{2i\left(\delta_0 - \frac{2\tau}{R}\right)\{\zeta(y)\}^2}{\{\zeta(y)\,\eta'(y_1) - \zeta'(y)\,\eta(y_1)\}^2}.$$

Case 1. $[\Re(x^{\frac{1}{2}}\varepsilon) \gg 1]$

Now

$$k = \frac{\zeta'(y_1)\,\zeta(y) - \zeta'(y)\,\zeta(y_1)}{\zeta(y)\,\eta'(y_1) - \zeta'(y)\,\eta(y_1)}$$

and

$$\zeta_{\nu-\frac{1}{2}}(x_1)/\eta_{\nu-\frac{1}{2}}(x_1) = k.$$

Since $\Re(x^{\frac{1}{2}}\varepsilon)$ is large and $\Im(x^{\frac{1}{2}}\varepsilon)$ is of order unity, then, by (56)

$$\zeta_{\nu-\frac{1}{2}}(x_1)/\eta_{\nu-\frac{1}{2}}(x_1) \doteq -1.$$

Hence, approximately, $k \doteq -1$ and so

$$\zeta'(y)/\zeta(y) \doteq \Psi'(y_1)/\Psi(y_1).$$

Thus,

$$\frac{\partial k}{\partial s} = \frac{2i\left(\delta_0 - \frac{2\tau}{R}\right)}{\{\eta'(y_1)/\eta(y_1) - \zeta'(y)/\zeta(y)\}^2\{\eta(y_1)\}^2}$$

$$\doteq \frac{2i\left(\delta_0 - \frac{2\tau}{R}\right)}{\{\eta'(y_1)/\eta(y_1) - \psi'(y_1)/\psi(y_1)\}^2\{\eta(y_1)\}^2}$$

$$= -2i\left(\delta_0 - \frac{2\tau}{R}\right)\{\Psi_{\nu-\frac{1}{2}}(y_1)\}^2$$

since $\zeta_4'\zeta_4 - \zeta_4'\zeta_4 = i$.

Case 2. $s - x_1 = 0(x^{\frac{1}{2}})$

For $s - x_1$ comparable with $x^{\frac{1}{2}}$, we have (since $x^{\frac{1}{2}}\varrho \gg 1$), the approximations

$$\zeta(y) \sim \frac{1}{2^{\frac{1}{4}}}\left(\varrho - \varepsilon + \frac{\tau}{R}\right)^{-\frac{1}{4}} \exp\left\{-\frac{2}{3}\sqrt{2}\,i\,x\left(\varrho - \varepsilon + \frac{\tau}{R}\right)^{\frac{3}{2}} + \frac{i\pi}{4}\right\}$$

$$\zeta(y_1) \sim \frac{1}{2^{\frac{1}{4}}}\left(\varrho - \varepsilon + \frac{\delta_0}{2}\right)^{-\frac{1}{4}} \exp\left\{-\frac{2}{3}\sqrt{2}\,i\,x\left(\varrho - \varepsilon + \frac{\delta_0}{2}\right)^{\frac{3}{2}} + \frac{i\pi}{4}\right\}$$

with similar expressions for $\eta(y)$ and $\eta(y_1)$; hence

$$\frac{2i\left(\delta_0 - \frac{2\tau}{R}\right)\{\zeta(y)\}^2}{[\zeta(y)\,\eta'(y_1) - \zeta'(y)\,\eta(y_1)]^2} = \frac{2i\left(\delta_0 - \frac{2\tau}{R}\right)}{\{\eta(y_1)\}^2\{\eta'(y_1)/\eta(y_1) - \zeta'(y)/\zeta(y)\}^2}$$

$$\doteq \frac{\sqrt{2}\left(\delta_0 - \frac{2\tau}{R}\right)}{4(\varrho - \varepsilon)}\left(\varrho - \varepsilon + \frac{\delta_0}{2}\right)^{\frac{1}{2}} \exp\left\{-\frac{4}{3}\sqrt{2}\,i\,x\left(\varrho - \varepsilon + \frac{\delta_0}{2}\right)^{\frac{3}{2}}\right\}.$$

* For convenience in writing the suffix $s - \frac{1}{2}$ is omitted from the ζ, η functions.

VOL. 2 1952

But the approximate value of k is,

$$k \doteqdot \frac{-i\left(\delta_0 - \frac{2\tau}{R}\right)}{8(\varrho-\varepsilon)} \operatorname{Exp}\left\{-\frac{4}{3}\sqrt{2\,i\,x}\left(\varrho-\varepsilon+\frac{\delta_0}{2}\right)^{\frac{3}{2}}\right\}.$$

Therefore,

$$\frac{\partial k}{\partial s} \doteqdot 2\sqrt{2\,i\,k}\left(\varrho-\varepsilon+\frac{\delta_0}{2}\right)^{\frac{1}{2}}.$$

Thus, since $\zeta_{\nu-\frac{1}{2}}(x_1) = k\,\eta_{\nu-\frac{1}{2}}(x_1)$,

$$2i - \{\eta_{\nu-\frac{1}{2}}(x_1)\}^2\left(\frac{\partial k}{ds}\right)_{s=\nu} \doteqdot 2i - 2i\sqrt{2}\left(\varrho-\varepsilon+\frac{\delta_0}{2}\right)^{\frac{1}{2}}\eta_{\nu-\frac{1}{2}}(x_1)\,\zeta_{\nu-\frac{1}{2}}(x_1),$$

$$\therefore F_j = 1 - \sqrt{2}\left(\varrho-\varepsilon+\frac{\delta_0}{2}\right)^{\frac{1}{2}}\eta_{\nu-\frac{1}{2}}(x_1)\,\zeta_{\nu-\frac{1}{2}}(x_1).$$

APPENDIX III

Approximation to normal propagation

Limitations of space forbid inclusion of the analysis necessary to establish conditions for this. We therefore briefly state the results here.

The inversion has little effect on propagation if

(i) It is too thick. This causes the reflection of even the most powerfully reflected rays to become weak. In the case of the Parabolic profile, this happens if τ exceeds about $800\,\lambda/\sqrt{h}$ (λ, h in metres). For the CARDINGTON inversion this is equivalent to the condition,

$$\tau < 22\lambda \text{ metres}.$$

(ii) The wavelength is too large. This causes P to become small. If P is less than about 0·7 the value of ε approximates to the normal value ε_0 given by (73). In the language of BOOKER [1], the inversion now lies below the track widths of all the modes.

(iii) If the refractive index drop through the inversion $(\Delta\mu)_0$, is too small. For inversions at the same height as that discussed in the text, it is found that there is little effect on propagation when $(\Delta\mu)_0$ becomes as small as 10^{-7}.

(iv) If the inversion is of insufficient extent. See § 9.

APPENDIX IV

Ray theory interpretation of characteristic roots

The characteristic roots for which X_j is not large are the most important at the distances in which we are interested, since Y_j is then least. By (62), the roots with X_j negative correspond to rays reflected from the inversion at obliquities less than limiting, the obliquity decreasing as $-X_j$ increases. Since Y_j increases as $-X_j$ increases, it is understandable, from reflection theory, that the contributions of the "modes" become weaker* as $-X_j$ increases. When X_j is positive (say > 1), the contributions of the "modes" become rapidly negligible, because, besides the increase in Y_j with increasing X_j, the factor F_j^{-1} decreases with extreme rapidity as X_j increases

* By a "mode" we mean a term in the field series (103).

much beyond unity. In this case, the "modes" correspond to rays incident upon the inversion at obliquities exceeding the critical. Such rays never reach the earth's surface, their least distance from it increasing with X_j. It is therefore easy to understand why "modes" for which X_j much exceeds unity give negligible contribution to the field near the earth's surface at the distances considered in this paper.

References

[1] Booker, H. G. and Walkinshaw, W.; Met factors in Radio Propagation. Phys. Soc. Report 1946 80. [2] Pekeris, C. and Ament, W.; Phil. Mag. 1947 **38** 801. [3] Westwater, F. L.; J. Inst. Elect. Engrs 1946 **93** (III A, No. 1) 79. [4] Attwood, S. S.; Radio Wave Propagation (London), 150 (1949). [5] Beghian, L. E. and Northover, F. H.; Admiralty Signal Estab. Research Paper M. 733, 1945. [6] Megaw, E. C. S.; J. Inst. Elect. Engrs 1946 **93** (III A, No. 1) 79 et seq. [7] Sheppard, P. A.; Phys. Soc. Report 1946 69–70. [8] Saxton, J. A.; Proc. Inst. Elect. Engrs III, 1951 **98** 360. [9] Saxton, J. A.; Proc. Inst. Elect. Engrs III, 1951 **98** 370. [10] Pekeris, C.; Phys. Rev. 1946 **70** 518. [11] Watson, G. N.; Proc. Roy. Soc. A 1919 **95** 83–99. [12] Watson, G. N.; Theory of Bessel functions. Ch. 7, page 201, Cambridge, 1944 Ed.

VOL.
2
1952

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2 .pp. 130 to 140. Pergamon Press Ltd., London

Liquid particles in atmospheric haze

SYLVAN RUBIN

Stanford Research Institute, California

(Received 20 August 1951)

ABSTRACT

Relatively dense air pollution haze was sampled by means of simple modifications of the jet impactor and the electrostatic precipitator. Microscopic observation indicates that the visible pollution consists mainly of stable liquid aerosol particles. It is suggested that these droplets are similar in composition to the smaller condensation nuclei commonly found in industrial urban atmospheres.

INTRODUCTION

In the course of studying the nature of "smog" in the Los Angeles area, a great many observations were made on the aerosol component of the polluted air, using a number of different techniques. It was found that the smog haze consisted largely of liquid droplets of rather low volatility, in the general size range from less than 0·2 to over 10 μ in diameter. This is in striking parallelism to the results of investigations of "natural haze" [1], [2], [3]. These studies have shown that the natural haze consists chiefly of small droplets, whether it occurs at sea level or at high altitudes, whether at high or low relative humidity, when observed in regions far from industrial sources of pollution. The conclusions regarding droplets in natural hazes were based on measurements of the optical properties of the atmosphere, and on studies of the haze particles collected by extremely fine fibres [2].

Two of the aerosol collection devices used to sample air for the present study of smog, obtained samples which were adapted to study under the microscope. These two methods were: (1) A modification of the jet impactor [4], and (2) an alternating potential electrostatic precipitator. Examination of the particles collected by these two means from industrial haze showed the particles to be liquid.

Both of these devices were very simple. The impactor collected aerosol samples in a small area on a glass slide; the electrostatic precipitator collected at a lower rate on a larger area of glass slide, making it possible to recognize individual particles. Both devices indicated a very wide variation in the concentration of the aerosols. In parallel runs, the two methods showed good qualitative agreement.

APPARATUS

A jet impactor collects aerosol particles by impinging a high velocity air stream on a surface in such a way that the air stream is required to turn in a very small radius of curvature. Small particles in the air stream do not have their momentum changed enough in a short distance to avoid striking the deflecting surface. Air enters through the jet tube and will reach sonic velocity at the jet if the impactor is operated with one atmosphere upstream and a half atmosphere or less downstream of the jet.

It was found unnecessary to use a "sticky" coating to collect the small aerosol particles from smog, as they adhered quite strongly to clean glass. Tests showed that the collection efficiency for particles smaller than 1 μ was essentially the same on dry glass as it was on a slide coated with a thin film of a glycerin gel.

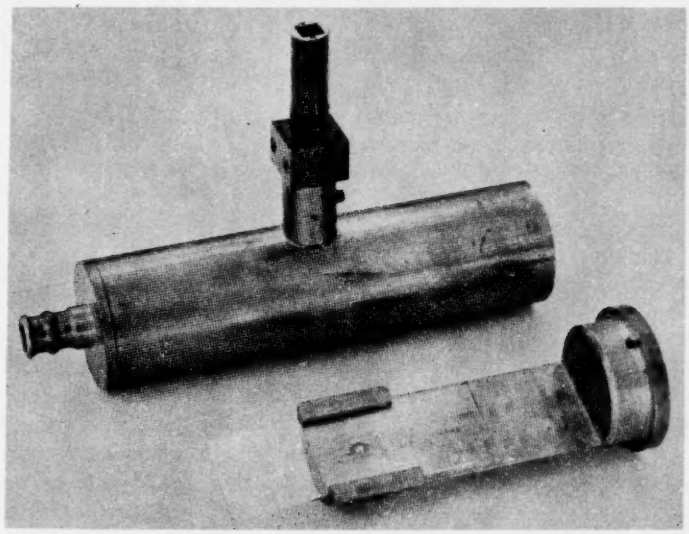


Fig. 1. Jet impactor.

The type of jet impactor which was chiefly used in this work is shown in Figs. 1 and 2. It consists of a brass tube about $1^1/_2$ in. in diameter, and about 5 in. long, open at one end and fitted with a hose connection at the other end. A short, heavy-walled sleeve is soldered into the side of the tube. The brass jet tube fits closely in this sleeve and is clamped in position by a set screw. The jet tube tapers at one end to a line opening about $^3/_8$ in. long and 0·004 in. wide. The tube is made in two pieces, separable along a plane including the axis of the tube and the long dimension of the jet, so that it may be disassembled for cleaning. The tube is held together by two screws threaded into external projections near the middle. The sample-collecting surface is a glass microscope slide, mounted on a holder inserted into the open end of the tube, thus sealing that opening. The slide fits into two slots and is held in place by a spring-loaded steel ball which presses against its back.

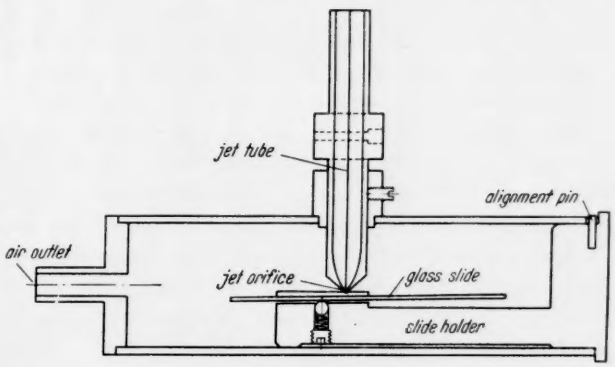


Fig. 2. Jet impactor.

As the drawing shows, the slide is held at a small angle to the impactor axis to facilitate insertion and removal without touching the end of the jet tube. The slide holder is aligned with respect to the jet axis by a pin which fits a notch in the edge of the housing tube.

The spacing between the jet and the surface of the slide is adjusted by loosening the set screw and sliding the jet tube in its holding sleeve. The gap between the jet and the slide is usually about 0·01 in. wide, so as to be comparable with the jet width.

This device is very convenient and effective as a detector of particulate matter and as an approximate indicator of the relative concentration of the aerosol. With a sample of about 1 ft³ of air—a size of sample satisfactory even in very dense smog—it is able to distinguish easily between very clear, open atmosphere and well-filtered air.

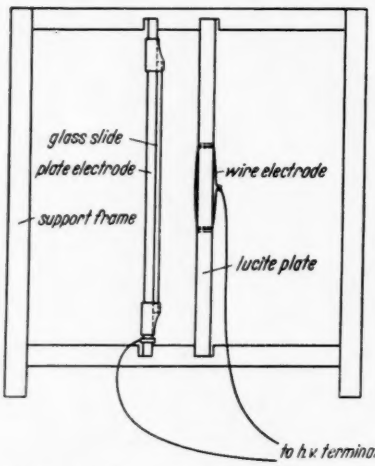


Fig. 3. Electrostatic precipitator.

The jet acts as a limiting orifice, giving a flow of approximately 0·2 ft³/min when the impactor is connected to a vacuum pump of sufficient capacity to induce sonic velocity. At this flow rate, a small laboratory pump is adequate.

The electrostatic precipitator was constructed in a very simple manner, as shown in Fig. 3. The 1×3 in. glass microscope slide was located in an alternating field produced by a *Ford* spark coil between a metal plate and a parallel wire electrode about a centimeter removed from the plate. The slide was taped to the plate electrode, and the wire electrode was held in place by threading the ends through a sheet of lucite. The electrodes were mounted loosely on an open wooden frame, so that air currents could freely circulate between them. A continuous corona discharge took place between the wire electrode and the surface of the slide when the device was in operation.

The purpose of this device was to collect samples of atmospheric aerosols as discrete particles in a form suitable for observation under the microscope. This arrangement operated very satisfactorily for this purpose. In dense smog (visibility about 2 miles), suitable samples are obtained within a few minutes.

Slides examined by dark-field microscopy after a run showed a maximum concentration of particles collected on the two areas of the slide opposite the ends of the wire electrode. The quantities collected diminished greatly toward the edges of the slide.

This electric collector was especially suited to investigating the upper range of haze concentrations, since the concentration of particles collected from clear air was so small that it was difficult to make the slides initially clean enough to note a significant increase in particle loading. However, by using special care in cleaning and handling slides, the device could be made effective for sampling aerosols in "clear" as well as hazy atmospheres.

Sampling

The experiments described here were all done at ground level in the Los Angeles area, mostly at Pasadena.

Impactor samples from the open atmosphere were generally taken in runs varying in length from a few minutes to a few hours, depending on the purpose of the sample and the clearness of the atmosphere. A run of 2 to 10 min is sufficient for correlating observations of the slide with atmospheric conditions. Enough material is collected from 0·4 to 2 ft^3 of air to be readily examined at low magnification, and to provide a representative sample of the atmospheric aerosol. Longer runs were used when the material collected was intended for chemical or physical analysis.

In sampling open atmosphere the only precaution taken in locating the impactor was to prevent its inlet from approaching surfaces from which dust might be blown. Most of the sample were taken 3 or 4 ft from ground level, with the apparatus on a platform or table.

The electric collector was operated in a similar location. Air was circulated past the electrodes by the natural flow of air. The collector was usually placed in a tray, or on a clean sheet of paper, to prevent any local deposits of dust on the platform from being blown into the device. Runs usually lasted $^1/_2$–1 hr, although some longer runs were made to obtain slides with a large assortment of particles in a single photomicrograph.

VOL.
2
1952

Microscope Techniques

After a sample had been obtained on a slide by either the impactor or the electric precipitator, it was examined with a microscope to determine the character of the collection. It is easily understood that the detection and recognition of liquid material on the slides depends markedly on the type of microscope illumination used.

Four types of illumination have been used in these observations, and some of their characteristics are discussed below. These four types were: (1) bright field, (2) dark field, (3) oblique lighting, and (4) phase contrast.

With bright field illumination, using the standard *Abbé* condenser and with all optical components axially centered, the liquid material on most of the sample slides was generally found to be either invisible or extremely difficult to see. All the early observations were made with a bright field condenser used in the conventional way, and it was some time before it was observed that liquid had been collected at all. On the electric precipitator slides, the bright field illumination shows only a relatively few particles on the slide.

When the same slide is examined by dark field illumination, there is an enormous increase in the concentration of visible particles. Some of the larger particles, not visible at all by bright field illumination, can be seen to be liquid droplets, but all the smaller particles appear only as points of light and have no recognizable structure. An impactor slide, when examined by dark field illumination, does not show liquid because it usually is not distributed in discrete droplets, and the large smears of liquid with low surface curvature do not produce contrast.

An oblique illumination can be produced by several methods. Obstructing one side of a bright field condenser, or decentering the mirror below the condenser, will produce an asymmetric illumination. A very convenient method found to have some advantages, although effective only with objectives of low numerical aperture,

was to lower a paraboloidal dark field condenser away from the slide until a bright field is seen, then back toward the slide until the field just begins to darken. Under these conditions, transparent substances having a non-planar upper surface become visible by a differential refraction effect. This arrangement is convenient because three types of illumination are available merely by varying the position of the condenser with the height control knob. Of course, this is effective only at low magnification and only moderate illumination intensity can be obtained. However, the obtainable illumination is usually adequate, and the wide variety of contrasts makes possible the recognition of various types of substances. The effect of the position of the condenser is shown by Figs. 7a and b.

The use of oblique illumination made it possible to identify most of the material collected by the impactor from very hazy atmosphere as a liquid which evaporated only very slowly at room temperature.

The particles collected by the electric precipitator are also readily visible under a phase contrast microscope. With this type of microscope, it is possible to satisfy the aperture conditions for high resolution, and still obtain sufficient contrast. It was possible to observe that the particle sizes of a sample collected from smoggy air ranged down to 0·2 μ in diameter and less. However, as in the case of dark field illumination, there is no visible difference in the appearance of solid and liquid particles, which one can see with the oblique illumination.

The oblique illumination makes use of the refraction of the light at the curved surface of the droplet of liquid to give contrast between it and its surroundings. Consequently, it is possible to recognize the hemispherical shape of the surface for particles as small as 0·5 μ in diameter, and thus distinguish individual droplets from solids.

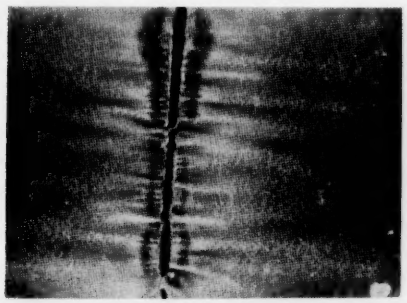
A statistical differentiation between liquid and solid particles may also be obtained by observing what fraction of the particles are not visible with bright field illumination, in comparison to the total number visible with dark field illumination.

OBSERVATIONS

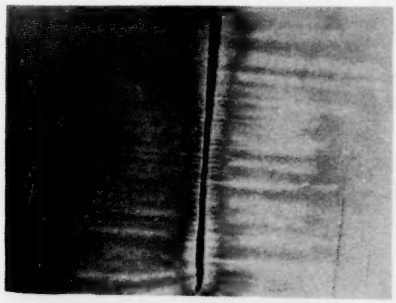
Impactor samples have been taken in the Los Angeles area at various times for more than two years since the liquid matter was first observed on these slides. The samples represent a considerable variety of atmospheric conditions, ranging from the very clean air present following a storm, to dense fog and strong smog. The amount of liquid found on a slide for a given volume of air passed through the impactor varied by a large factor, but always in conformity with the haziness of the atmosphere. Solid materials were also collected on the slides, and the proportion of liquid to solid also varied considerably, in a way which can be related to the type of haze and the relative humidity in the atmosphere.

The aerosol particles are collected by the impactor on the glass slide in a very narrow band just opposite the jet opening. If much liquid is collected, the air stream forces some of the accumulated liquid to flow along the slide away from this central line, forming a pattern of multiple parallel streaks at right angles to it. This pattern can easily be seen on the slide with the unaided eye when side-lighted against a relatively dark background. The existence of such a pattern is a very sensitive test for even the slightest "fogginess" in the air. The length of these

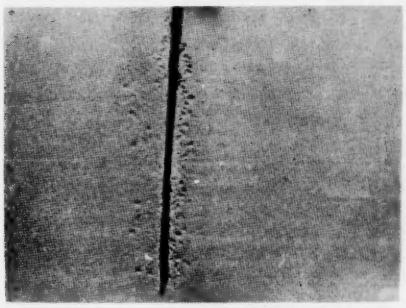
lateral streaks may be a centimetre or more in foggy weather. In dryer and clearer air they are shorter, and they do not appear at all from air which is either very clear or very dry.



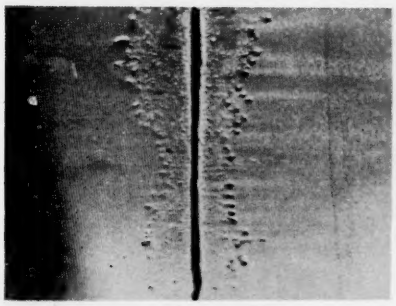
Dense smog. 11.30 a.m. 7 Nov.



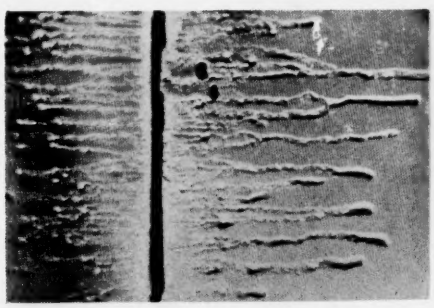
Moderate smog, 1.45 p.m. 7 Nov.



Moderate smog, 3.15 p.m. 7 Nov.



Light morning fog, 10 a.m. 8 Nov.



Very dense smog, 11.15 a.m. 8 Nov.

Fig. 4. Series of photomicrographs showing variations in relative amount of liquid collected during 15 min impactor runs at Pasadena during fog and smog in November, 1950.

Under the microscope, using oblique illumination, these lateral streaks are easily detected and examined. They usually end in a relatively large droplet of the liquid. Most of the solid aerosol particles remain in the central band where they are initially collected, although a small fraction may be carried along with the liquid. A series of samples taken on November 7–8, 1950, during foggy and smoggy weather show this very well. These slides are shown in Fig. 4.

VOL.
2
1952

The simplest technique for distinguishing liquid from solid particles on the microscope slide is probing with a very sharp needle. We have made such probe needles from dental probes, by cutting off the barbed end, and sharpening the tip of the needle to a very acute angle. The point is finished by careful polishing on crocus cloth, while spinning the needle in the fingers. Points of the order of a few microns across the tip can be made in this way. For probing particles, the needle is best used in a micromanipulator. It is possible by this means to obtain a relative estimate of the viscosity of liquid matter on the microscope slide, by observing how rapidly the liquid fills in a line drawn across it.

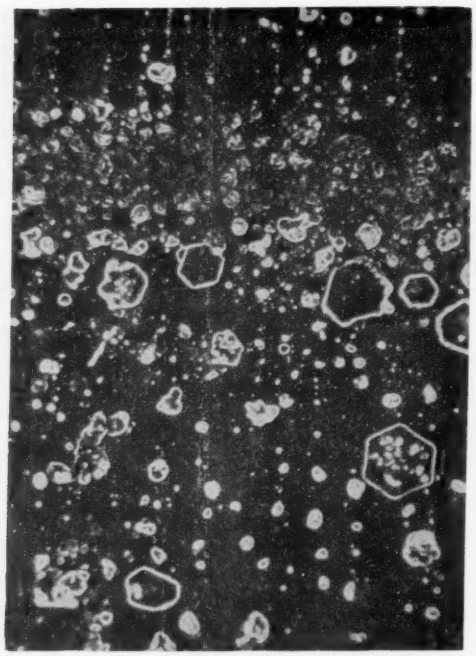


Fig. 5. Crystals formed in impactor sample of fog.

The liquid material was found to be only slightly volatile as compared with water. Droplets of pure water of the same size as the droplets of liquid found on the slides will evaporate in a few seconds at a relative humidity of about 50 %, whereas the liquid droplets collected from the atmosphere will persist in storage at this humidity for several weeks.

This liquid is collected in the Los Angeles area most abundantly from fogs. In such samples it is accompanied by a small proportion of solid matter, most of which appears to be very finely divided, amorphous black material. Very little crystalline material appears in the fresh sample, but after a few hours, many crystals can be found, increasing in number and variety as the liquid evaporates. An example of such crystals is shown in the photomicrograph of Fig. 5.

When smog is present, the liquid is again the most abundant component of the aerosol collected, but it is generally collected in smaller quantities than from fog, and is also more viscous. Furthermore, the black amorphous material found in fog is less abundant and, instead, there is more transparent crystalline material in the central streak.

As the liquid gradually evaporates, increasing numbers of crystals can be found along the lateral streaks. These crystals generally appear as very thin, hexagonal, colourless, optically isotropic platelets. A rather large crystal of this type usually can be found where there was a droplet at the end of the streak. The droplets themselves show a wrinkled surface after they have evaporated for several days, indicating that some transparent resinous material may be being precipitated at their surface. The liquid becomes very viscous as it evaporates, whether the evaporation is slow or made rapid by heating.

The liquid darkens on heating, first to a straw colour at a temperature below 100° C, then to a dark caramel at 200–300° C. It can be observed to boil if the slide

is heated directly with a flame, but the boiling point has not been measured. Since it evidently contains much dissolved matter, the boiling point presumably depends on how much evaporation occurs before boiling, and therefore on the rate of heating, and the surface to volume ratio.

In clear weather, especially if the humidity is low, there is very little material collected. Most of that which is found is transparent, crystalline matter, and only a small fraction is liquid. The proportion of liquid found under these conditions depends on the relative humidity. Small quantities of the liquid have been detected in samples taken at a relative humidity as low as 20%.

Fig. 6. Impactor sample of gasoline engine exhaust.

It is significant that impactor samples collected from the combustion gases of the common fuels invariably contain a great deal of liquid with characteristics similar to those just described. In comparison, the open atmosphere is relatively free of the liquid. Fig. 6 and 7 show slides obtained from a gasoline engine exhaust and from a power plant stack.

Micrurgic analyses of samples collected by impaction from the Los Angeles atmosphere show the presence of a wide variety of substances [5]. It is interesting that various investigators find that different substances predominate in the atmospheric aerosols, depending on the regions in which their studies were made [2], [6], [7], [8]. Sodium chloride, sulfuric acid, nitrous and nitric acid, ammonium compounds, organic substances, and others are among the substances reported as constituents of condensation nuclei in rural and urban regions.

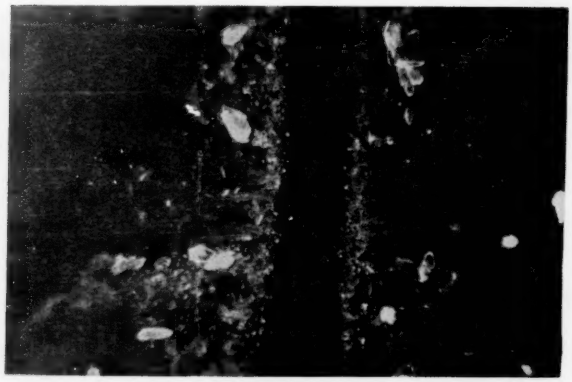
Only a few experiments were made with the electric collector. The purpose of the runs was to provide an independent means of collecting particles, as a check on the observations made with the impactor, and to show the collected material as discrete particles, without agglomeration. Following preliminary tests to determine operating characteristics, a number of runs were made in parallel with the impactor in various atmospheric hazes.

These parallel runs showed that the particles collected by the electric collector correlated well with the samples obtained by the impactor. They gave some

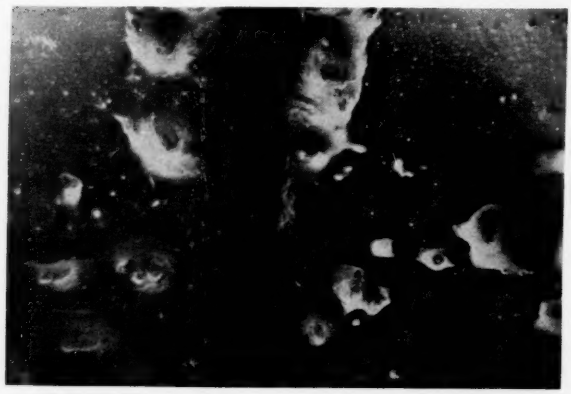
VOL. 2 1952

indication of the range of particle sizes in the haze, and also showed that nearly all the particles collected from the denser smog hazes were liquid. Fig. 8 is a photomicrograph of a portion of one such slide, showing the variation in droplet size.

A very large number of particles were collected in foggy or smoggy weather as compared with the quantity collected in clear weather. Although there was no control of the air flow past the electrodes, wind velocities were quite small during all these tests. It was assumed that the concentration of particles collected would be approximately proportional to the concentration in the atmosphere at the time of exposure. Although qualitative, the comparison with the haziness of the atmosphere, and with impactor samples, indicated that this assumption was approximately correct.



(a) With paraboloidal condenser set for dark field.



(b) With paraboloidal condenser set for oblique lighting.

Figs. 7a and b. Impactor sample of power plant stack gas.

Some samples taken in smog with a *Casella* thermal precipitator yielded particle concentration in the range of several thousand per cm^3, when the samples were counted under dark field illumination. The thermal precipitator was operated at a very slow flow rate, which gave approximately 100 % collection. The thermal precipitator was modified by replacing the original hot wire with a $^1/_{16}$ in. wide tantalum strip, so that the achievement of 100 % efficiency could be observed as a reduction in the width of the collection area.

The particles could be seen much more readily on the slides from the electric collector, especially by oblique microscope illumination, when the slides had been previously treated with General Electric *Drifilm*, which makes the glass surface hydrophobic. The use of *Drifilm* increased the contact angle between the droplets and the glass, thus enhancing the refractive effect of the droplets. It also made all the droplets circular in cross-section, whereas the larger droplets on untreated glass slides were frequently irregular in shape, due to partial spreading out on the glass. The *Drifilm* had a similar effect on the appearance of the liquid collected on impactor slides.

By microscopic examination of slides placed in a closed humid stage, CADLE [9] has shown that the material collected on impactor slides is very hygroscopic. He has also demonstrated that the collected liquid is miscible with water, but not with xylene.

The hygroscopicity of the liquid collected, and its behaviour on hydrophobic surfaces, as compared with oil droplets, indicates that the liquid aerosol collected from smog or from fog is mainly aqueous in composition. It evidently reaches equilibrium with water vapour at relative humidities considerably below 100%, thus making possible the occurrence of stable fogs at low humidity, on occasions when the smog aerosol is available in sufficient concentration.

There does not appear to be any well defined transition between the liquid collected from an unmistakable fog, and that collected from the haze at relative humidities below 100%. In fact, the relative humidity is usually somewhat below 100%, even when the haze is obviously a dense fog [10]. This is similar to the observations of BRAZIER in Paris, who reports that from 1919 to 1933 the relative frequency of occurrence of fogs at relative humidities below 100% gradually increased.

Fig. 8. Photomicrograph of electrostatic precipitator sample of smog. The large droplet near the centre is about 6 μ in diameter.

CONCLUSIONS

The fact that both natural and industrial hazes seem to consist largely of a liquid aerosol indicates that this type of aerosol can be readily formed in the atmosphere under various conditions.

The differences in the dissolved material in the industrial smog aerosol and that in the natural haze aerosol droplets indicates that the liquid aerosol is not formed from any unique substance, but may be produced from a variety of compounds available in sufficient concentrations. This is reflected in the relative complexity of the aerosol collected from an industrial atmosphere as compared with that collected in non-industrialized areas.

It is known that combustion processes contribute large quantities of condensation nuclei to the atmosphere. The large quantities of hygroscopic liquid found in samples of combustion gases, and the high concentration of the liquid aerosol found in foggy atmospheres, tend to support the hypothesis that the liquid aerosol which constitutes the visible haze acts as condensation nuclei for fog formation.

Various investigators have indicated that the mean effective size of condensation nuclei is below a diameter of 0·1 μ. This is not necessarily in contradiction to the conclusion stated here. It is known that the supersaturation factor for vapour

VOL. 2 1952

condensation decreases with increasing diameter of condensation nucleus [11], and it may be assumed that combustion nuclei usually occur with a size distribution decreasing roughly exponentially with diameter. On this basis, the effective size of the nuclei, as measured by condensation experiments, would be the largest diameter for which there is a sufficient concentration of nuclei for the amount of condensation, and the effective average nucleus size would therefore be a function of the concentration of condensation nuclei. AITKEN has shown that only a few hundred nuclei per cm^3 in the atmosphere are sufficient for ordinary fog formation. When the condensation nuclei from combustion processes are in sufficiently high concentration, as in smog, then the same particles which produce the visible haze also may be the major part of the condensation nuclei.

Because of the very rapid variation of the optical scattering cross-section of dielectric particles with particle diameter, for diameters below the wavelength of visible light, the maximum "haze" effect for a given mass concentration of spherical particles of refractive index 1·33 occurs at about 0·7 μ diameter. If we assume a homogenous aerosol of particles of 0·7 μ diameter, of refractive index 1·33, and at a concentration of 100 particles per cm^3, the theoretical "visibility"* will be 20 miles** (for other concentrations, the visibility is inversely proportional to the concentration). This calculation indicates that at concentrations at which an aerosol forms a visible haze, there will be enough particles in the 0·1 to 1·0 μ size range to provide all the necessary nuclei for fog formation in that haze.

The frequent occurrence of low humidity fogs in the Los Angeles area, and their observation in other large urban centres appears to be a consequence of condensation on these large liquid aerosol particles.

Acknowledgements—This paper comprises part of a research programme on the nature of smog. This work is being done at *Stanford Research Institute* under the sponsorship of the *Western Oil and Gas Association.* The author is indebted to Mr. PAUL L. MAGILL for his encouragement and assistance in the preparation of this paper, and to Mr. R. W. BENOLIEL and Dr. R. D. CADLE for helpful discussions.

REFERENCES

[1] DESSENS, H.; Ann. Géophys. 1946 **2** (1) 68–85. [2] DESSENS, H.; Ann. Géophys. 1947 **3** (1) 68–86. [3] MOLLER, F.; Met. Rdsch. 1948 **1** 212–215. [4] MAY, K. R.; J. Sci. Instr. 1945 **22** 187–195. [5] CADLE, R. D., RUBIN, S., GLASSBROOK, C. I., MAGILL P. L.; Arch. Ind. Hyg. and Occ. Med. 1950 **2** 698–715. [6] CAUER, H. and CAUER, G.; Balneologe 1941 **8** 345–353. [7] COSTE, J. H. and WRIGHT, H. L.; Phil. Mag. 1935 **20** 209–234. [8] SALLES, E.; Rev. Sci. (Paris) 1940 **78** 77–84. [9] Private communication. [10] NEIBURGER, M. and WURTELE, M. O.; Chem. Rev. 1949 **44** (2) 321–335. [11] LANDSBERG, H.; Ergeb. Kosm. Physik 1938 **3** 155–252.

* Assuming that the visual limit of detectable contrast = 0·02.

** In the Los Angeles area, this visual range is visible as a haze due to the obscuration of nearby mountains.

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 141 to 154. Pergamon Press Ltd., London

Measurement of winds in the ionosphere

G. J. PHILLIPS

Cavendish Laboratory, Cambridge

(Received 23 September 1951)

ABSTRACT

The fading of an ionospheric echo at several receiving points on the ground has been used to measure the velocity of horizontal movement in the ionosphere. A systematic series of observations has been made for a period of over two years, on a frequency of 2·4 Mc/sec. Reasons are given for supposing the results to refer to the motion of the air at heights in the range 100–120 km. Regular daily and seasonal changes of direction were found, partly attributable to a solar tide having an amplitude of 16 m/sec. The observed phase of this tidal motion is not consistent with the simple dynamo theory of geomagnetic quiet-day variations. Remaining changes of a less regular nature suggest that rapid variations of wind at a fixed level, or differences in wind at closely separated levels, may occur.

VOL. 2 1952

1. INTRODUCTION

PAWSEY [1] first described a method by which wind-like movements in the ionosphere can be studied from a single observing station by the use of quite simple apparatus. More recently MITRA [2] and KRAUTKRÄMER [3] have described measurements made by this method. It is the purpose of this paper to describe the results of a longer series of measurements of this kind made on several days in each month from January 1949 to June 1951.

The observations over this period have shown a considerable amount of regularity which appears to be of significance in the study of the upper atmosphere. It is understood that much of this regularity is also observed at Washington, D.C. and Montreal, Canada, where similar experiments are being carried out.

MITRA used a pulsed radio transmitter, and three receivers at the corners of a triangle with sides of about one wavelength (100 m). Simultaneous records were made of the amplitude of the echo pulse returned to each receiver from an ionospheric layer. When these records were studied it was found that the fading of the echo was often similar in nature at the three points but with a time-lag superimposed. On these occasions it was assumed that the ionosphere had irregularities of electron density so that the reflected waves formed an irregular diffraction pattern on the ground, and, further, that as the irregularities moved horizontally, they caused the pattern to move along with them in the same direction but with twice their velocity. The time delays between the fading curves for the three receivers could then be used to deduce the magnitude and direction of the velocity of the ionospheric movement.

There are two major difficulties in the interpretation of records of this type. The first arises when the records for the three receiving points are so dissimilar that they cannot be associated with a diffraction pattern which simply moves horizontally. The pattern must then be assumed to change in form as it moves. BRIGGS, PHILLIPS and SHINN [4] have considered this case in some detail and have shown that inaccuracies can result if the random changes are ignored. The present results are therefore examined from this point of view in Section 3.

The other difficulty arises when an attempt is made to specify the atmospheric level at which the wind-like movements occur. A movement of the radio diffraction pattern over the ground could be produced by a movement of an irregular portion of the ionosphere at any level up to and including the point of reflection. In general there is no method of locating the relevant level but some indirect evidence concerning the height of the observed movement is given in Section 5.

2. EXPERIMENTAL ARRANGEMENTS

A pulsed sender emitting 50 pulses per second was used in conjunction with four receiving aerials situated as shown in Fig. 1a. Each receiving aerial was a single-turn shielded loop coupled to a coaxial cable by a shielded transformer. The aerial cables were laid along the ground to a central point. Here they were switched in turn to a single receiver

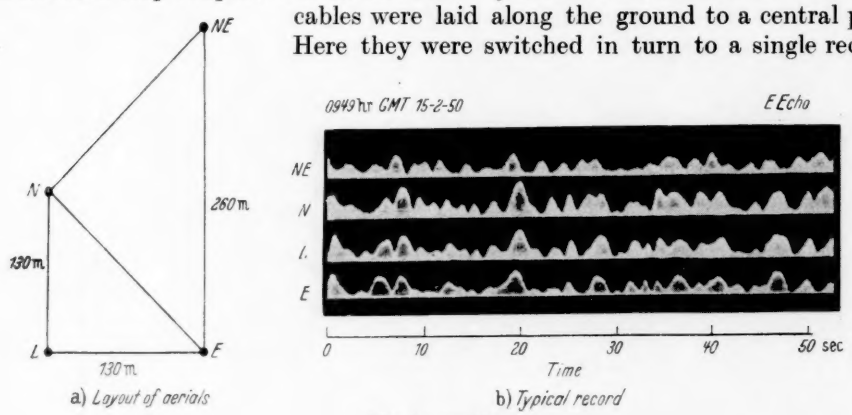


Fig. 1 a and b.

by high-speed relays, which were synchronized with the radio pulses in such a way that only one pulse (with ionospheric echoes) was received each time an aerial was connected to the receiver. An electrical gate selected a single echo reflected with a given delay-time from the ionosphere, and the output of the receiver was applied to a cathode-ray oscillograph so that the amplitude of this echo was represented by the length of a bright line on the screen. The spot on the screen was shifted to different positions according to which receiving aerial was being used. As a result four traces were produced of lengths proportional to the amplitude of the echo at each aerial. These were photographed on a moving film to give the type of record shown in Fig. 1b. Periods of recording were kept short in order to conserve film.

Only three aerials at the corners of a right-angled triangle (usually L, N and E of Fig. 1a) were required for calculating the results, but a fourth aerial was available as a check.

3. ANALYSIS OF THE RECORDS

The film records were viewed by projection from an enlarger. It was found that for most of them the time shift between similar fluctuations at the spaced aerials could be estimated by eye with the aid of a series of sliding strips. Each strip was provided with identical scale marks, and had one of the four amplitude records projected upon it; the strips could then be shifted to correspond to the time delays.

An ionospheric velocity v in a direction making an angle Φ with north would be expected to produce a movement of the diffraction pattern on the ground in the same direction with velocity $2v$. The time displacements between the fading curves recorded on pairs of aerials separated by a distance d in the N–S and E–W directions would be $d\cos\Phi/2v$ and $d\sin\Phi/2v$ if there were no complication caused by changes of form in the pattern as it moved. The records were first analysed on the assumption that there were no such changes, and the following procedure was then used to test the truth of this assumption. For each record an estimate was made of the speed of fading and also of the mean spatial gradient of the amplitude, R, in the radio diffraction pattern on the ground. The first quantity was found by counting the number of maxima n occurring in a fixed time t. The second quantity was obtained from the instantaneous difference between the amplitudes recorded at two adjacent receiving aerials (spaced at distance d); the average magnitude of the difference was expressed as a fraction $\Delta/\bar{R}$ of the mean amplitude.

These measurements may be related to the normalized temporal and spatial gradients of amplitude; thus

$$S = \frac{\overline{\left|\frac{\partial R}{\partial t}\right|}}{\bar{R}} = 1{\cdot}6\,.\,\frac{n}{t},$$

$$G = \frac{\overline{\left|\frac{\partial R}{\partial x}\right|}}{\bar{R}} = \frac{1}{d}\cdot\frac{\Delta}{\bar{R}}.$$

VOL.
2
1952

The first relation is accurate only for fluctuations of a detected random signal having a radio frequency spectrum in the form of a narrow Gaussian curve, but it may be shown that it is not seriously in error in the case of narrow band spectra having somewhat different forms. For a rectangular band-pass spectrum for example the numerical factor is 1·8 (see Rice [5], Furth and MacDonald [6]). The second relation assumes that the aerial spacing, d, is small compared with the scale of the diffraction pattern. In the experiment the spacing (130 m) was small enough for the relation to hold fairly well.

It was shown by Briggs *et al.* [4] that if the diffraction pattern were moving steadily without change of form, the quantity S/G would represent the velocity of the pattern. They also show that if random changes occur neither this ratio nor the velocity apparent from the time delays represents the true velocity of the pattern; moreover one quantity is affected more than the other, so that they are no longer equal to one another in these circumstances. A test of the presence of random changes can therefore be made as follows.

The quantity $n\bar{R}/\Delta$, which is proportional to S/G, was plotted against the velocity derived from the time shifts. A diagram showing the resulting points for a typical month's results is given in Fig. 2. In the absence of random changes in the diffraction pattern, both coordinates of a point would have been a true measure of the velocity, and then all points would have been on a straight line through the origin. The gradient of this line may be shown to be $t/1{\cdot}6\,d$ from the relations above.

However, in practice there was a considerable amount of scatter of points about this line. This was due to two causes. Firstly the estimates of n and $\Delta/\bar{R}$

were made only by rapid visual inspection of the records, and were therefore only approximate. Secondly some records were presumably affected by random changes. It was nevertheless possible to reject those records likely to give serious error, and this was done if the corresponding point in the diagram did not lie between lines through the origin differing in slope from the ideal value by about $\pm 50\%$. It is in theory possible to correct a velocity for the effect of random changes if the quantities mentioned above are known accurately [4]; it was, however, sufficient in the present case not to attempt to make further use of the rejected records (about 20%) and to retain the velocities as calculated from the time delays for the remainder.

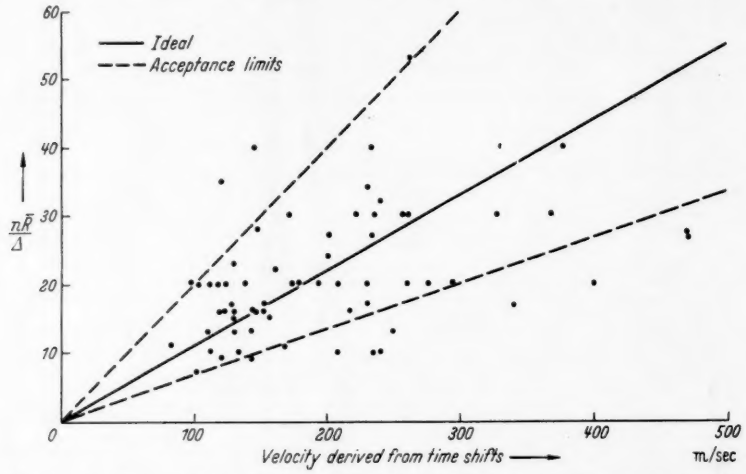


Fig. 2. Scatter plot for results in October 1950. The velocities here refer to the diffraction pattern on the ground.

4. RESULTS

Observations were made on three or four consecutive days in nearly every month of the period January 1949 to June 1951. Film records of 3 min duration were taken at about half-hour intervals throughout the day and night, using a radio frequency of 2·4 Mc/sec. On this frequency echoes were obtained from the E layer by day and from F or Sporadic E (E_s) layers by night. For about a third of the night-time results echoes from the E_s layer were used, the proportion being larger in summer than in winter.

On many occasions observations could be made in rapid succession on echoes from the F layer and the E or E_s layer, and it was then found that, even though individual records could give appreciably different results, the mean trend for the two layers was often very similar. It has already been pointed out that a wind-like movement in any irregular layer at or below the point of reflection could give rise to movements of the diffraction pattern at the ground, and the similarity of the results obtained with echoes from the E and F regions shows that either

(a) movements in E and F layers are themselves very similar, or

(b) echoes from the F layer are often affected by irregularities in the E layer sufficiently for the results to be characteristic of the E layer.

For this reason all the results are for the present taken together as applying to some height within Region E; further discussion of the height of the movements will be given later.

The magnitude and direction of horizontal motion in the ionosphere calculated from each reliable record was resolved into its components in the N-S and E–W directions. These were then plotted separately against time of day using all the readings for one particular month. Fig. 3 gives the results for October 1950 as an example. It can be seen that there was a considerable spread of values, but the mean trend of each component appeared to be much the same from day to day and to

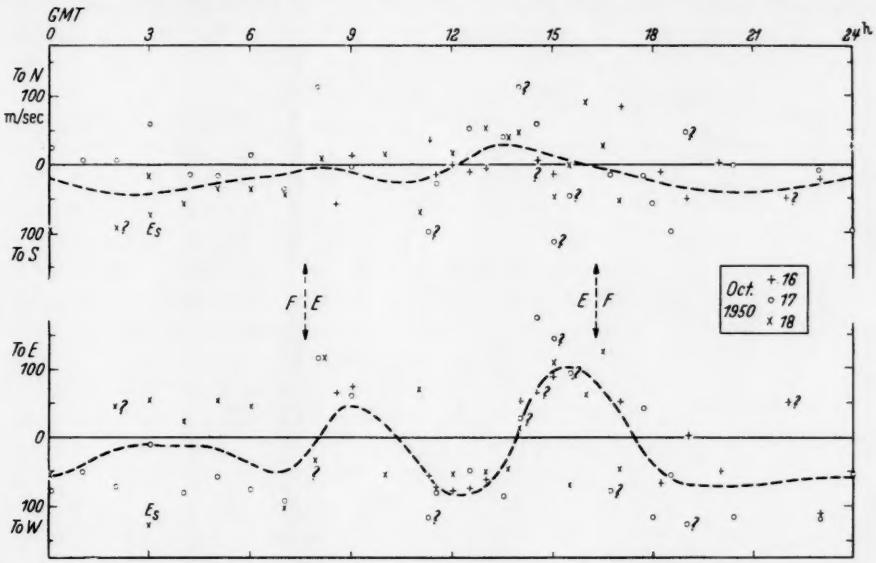


Fig. 3. Components of ionospheric velocity observed in October 1950.

show interesting variations throughout the 24 hrs. A smooth curve was drawn to represent this average variation for each month. These curves were then used to produce the polar plots given in Fig. 4, which show in a concise form the variation of both magnitude and direction of the smoothed horizontal velocity vector throughout the day.

The histograms of Fig. 5 show how the magnitudes of the horizontal velocities were distributed; (a) includes all observations, (b) includes only those referring to the three summer months, and (c) only those for the three winter months, throughout two years. They show that larger velocities tend to occur in winter.

On the occasion of two severe magnetic storms interesting records were obtained. An example is given in Fig. 6. Very fast fading occurred, although the differences between the records at any instant were not appreciably greater than usual. We conclude that the scale of the diffraction pattern on the ground was not greatly changed, and that the faster fading was mainly due to increased velocities in the ionosphere. On several storm records of the type shown in Fig. 6 there was a small but distinct time shift indicating a very fast movement towards the west. During

VOL. 2 1952

the storm of 28th October 1949 the velocity appeared to be greater than 500 m/sec at 0035 hrs., dropping to 250 m/sec in an hour. These values were deduced from

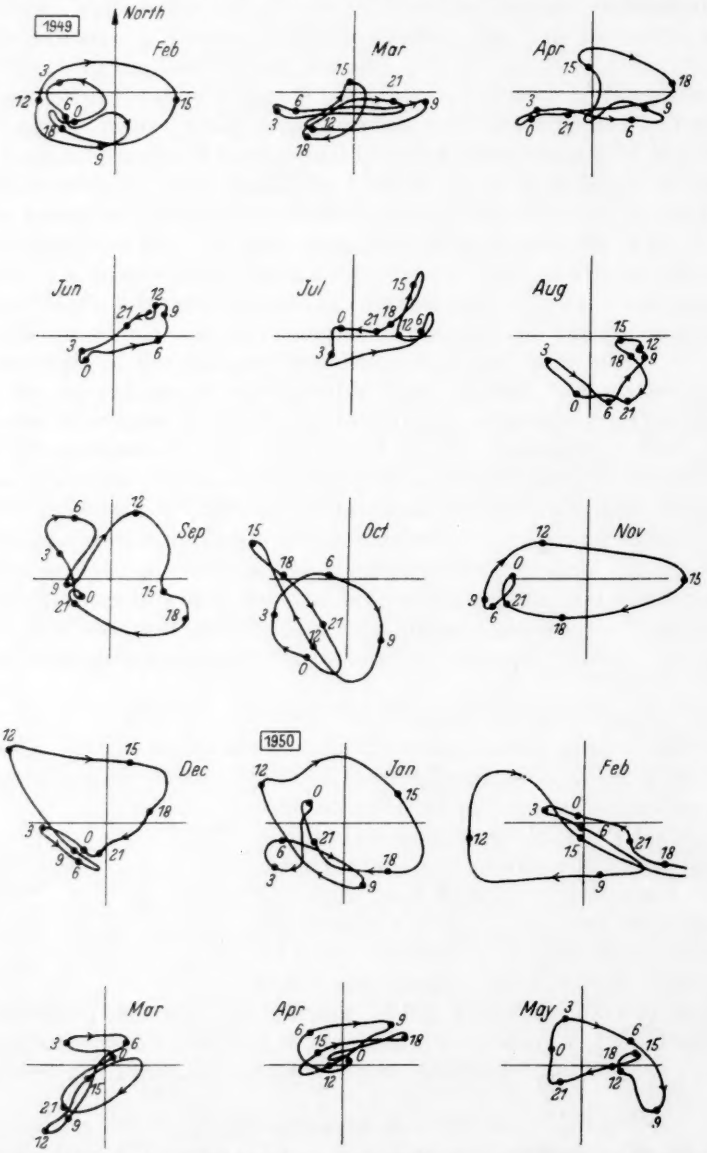
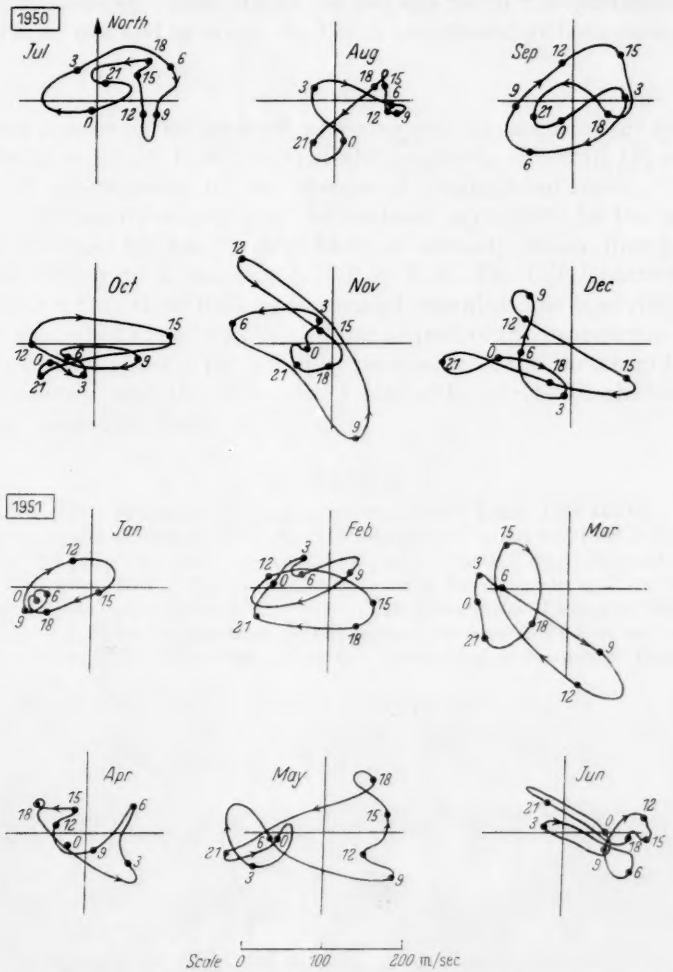


Fig. 4. Polar plots. These give the variation throughout 24 hrs as an average of the results for three
the origin to the appropriate point on the curve.

reflections from the *F* layer, and in this case probably refer to movements near the height of reflection at about 450 km. Exceptional observations such as these have been excluded from the general analysis.

We have fortunately been able to collaborate with SALZBERG and GREENSTONE [7] at the Bureau of Standards, Washington, D.C., U.S.A. and CHAPMAN [8] of McGill University, Canada, who are making measurements by the same method using very nearly the same radio frequency. Simultaneous observations have given results for the daily variation which show considerable similarity when referred to local time.

Because of this dependence on local time, the variations of velocity were examined for periodicities of 12 and 24 hrs. For this purpose a harmonic analysis was made of the N–S and E–W components separately. In Fig. 7 the 12 hr harmonics are shown



consecutive days. The direction towards which the wind is blowing is represented by a line from The figures denote the time in hours GMT.

for each month by a point whose distance from the origin indicates the magnitude of the harmonic, and whose angular position gives the time of its maximum, movements towards the north and east being considered positive.

VOL. 2 1952

The overall mean 12 hr harmonics are indicated by a cross. Circles of radius equal to the standard deviation of the mean are also shown. The means are seen to be statistically significant for both the N–S and E–W components, because in each case the origin lies well outside the circle.

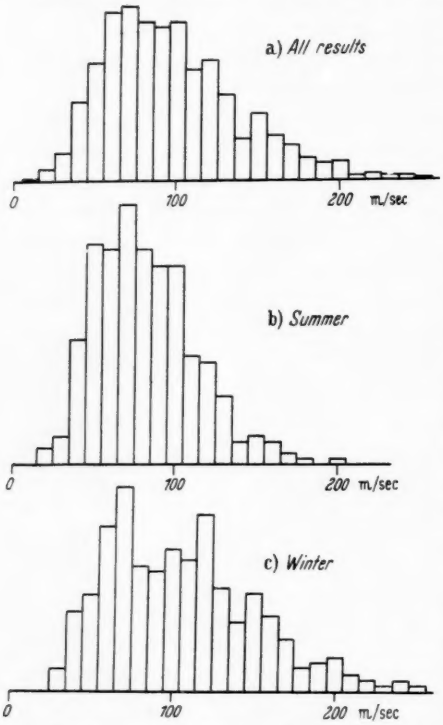


Fig. 5 a–c. Distribution of velocities.

The results for the 24 hr harmonic gave a significant variation for the E–W component only. This had an amplitude of 33 m/sec, the maximum velocity towards the east occurring at 1320 hrs.

Finally in Fig. 8 is given the mean overall horizontal drift in each month. The points in this diagram are such that a line from the origin to each point gives this quantity in vectorial form for the month indicated. The axes correspond to true N–S and E–W directions. An interesting seasonal change may be observed; the average drift is towards the East in summer and towards the West in winter.

5. Discussion

(i) *Height at which movements occur*—It is not easy to ascertain the height to which the measurements of wind velocities refer. Consider first the results obtained for reflections from the *E* layer. Fading of the echo could be caused by movements of irregularities of ionization at any height below the level of reflection down to a height where there is no appreciable refraction or absorption of the radio waves.

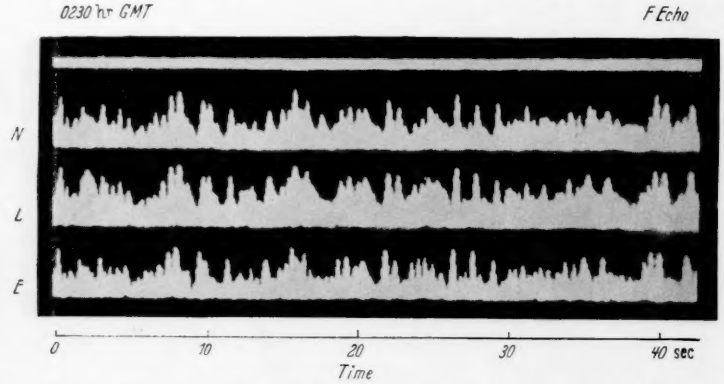


Fig. 6. Record taken on 28th October 1949 during a magnetic storm.

This means that the observed horizontal motion occurs between about 80 km and the height of reflection at 110–120 km. There is, however, some indirect evidence

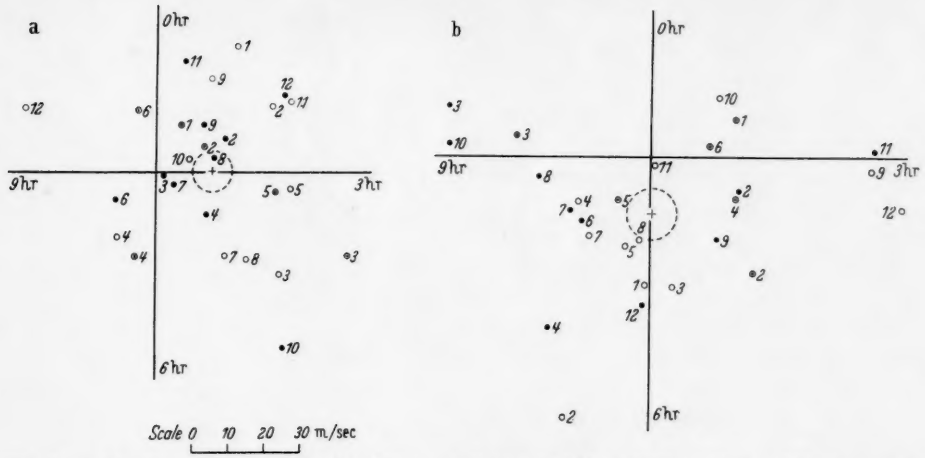


Fig. 7 a and b. Harmonic dials showing 12 hr periodicity. ● 1949; ○ 1950; ⊙ 1951; *a* 12 hr harmonic for N–S component; *b* 12 hr harmonic for E–W component. The amplitude and phase of the harmonic for each month is represented by a point. The type of point gives the year, and the month is indicated by appropriate numbers. The angular position of the point gives the time of maximum movement towards the north or towards the east. The mean harmonic is represented by a cross with a circle giving the standard deviation.

VOL. 2 1952

that the relevant height is near the point of reflection. Thus FINDLAY [9] has described how, on some occasions, he has observed isolated "clouds" of electrons usually just below the *E* layer and has been able to measure the magnitude (but not direction) of their horizontal movement. In FINDLAY's method the shortest distance of approach of the cloud was also measured. Since the clouds were only observable while nearly overhead, this corresponds very closely to the true height. Most of his clouds were at heights greater than 100 km and their speeds were distributed very much as shown in Fig. 5a. This is some indication that the present results also could very well refer to heights above 100 km. On one occasion observations made simultaneously by the two methods gave values for the velocity in good agreement. It was then fairly certain that the spaced receiver method applied to a height of 105 km, which was the height of the cloud observed on this occasion.

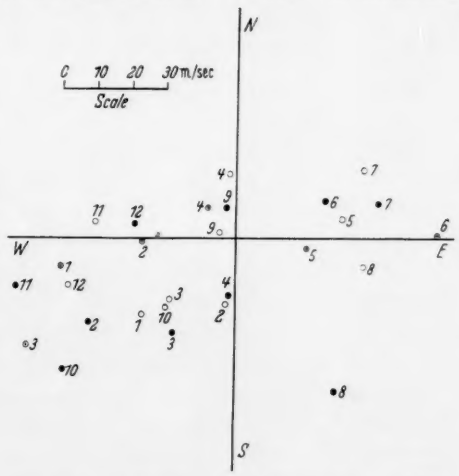


Fig. 8. Mean velocities. ● 1949; ○ 1950; ⊙ 1951. The mean velocity for each month is represented by a point labelled with the appropriate number. A line from the origin to the point gives the direction towards which the mean wind is blowing.

Again BRIGGS [10] has published records of the fading of an echo from Region *E* over a band of radio frequencies round about the frequency of 2·4 Mc/sec used in the present experiments. He sometimes found the fading more rapid above a

certain frequency, as in his Fig. 18 and 19. A wind observation on this more rapid fading would be largely controlled by irregularities near the point of reflection. For in this case the behaviour of the lower frequencies, which have passed through the whole of the region up to a slightly lower level, shows that the irregularities there are much less important. Under conditions such as these the measured wind would apply to a height very close to the height of reflection.

The fading of an echo from the *F* layer might be governed by the density and irregularity of the ionization in Region *E* as well as the irregularity in the *F* layer itself. It has already been mentioned that the somewhat similar average results obtained on the *F* and *E* layers might therefore be due to the predominant effect of *E* ionization or to the similar movements of the layers. It must be remembered however that under conditions when *E* and *F* echoes may be observed simultaneously, or within a short period, using the same working frequency, that frequency is never far from the penetration frequency for the *E* layer. It is then not unlikely that both types of echo will give results applicable to the height of the maximum of ionization in Region *E*. On the other hand when the *E* critical frequency is well below the working frequency it is expected that the measured movement would apply to the *F* layer.

Evidence that wind-like movements exist in *F* Region has been given by HEWISH [11]. He has shown that observations of the "scintillation" of radio-stars can be used to determine the velocity by a method similar in principle to the one described here. In general the magnitudes of his velocities are about equal to or slightly greater than those given here. A preliminary detailed comparison of his results for the E–W component with results for an echo from the *F* layer at the same time has suggested that on some occasions they agree in magnitude, and in form of variation, while at other times they are not correlated.

The following tentative conclusions are drawn regarding the height of the measured motion.

(a) In the daytime, when reflections employed were always from Region *E*, the results apply to some height between 100 and 120 km. Changes of measured velocity might occur because a different atmospheric level within this range had become operative, or because the velocity at the appropriate level had itself altered.

(b) In the night-time, results obtained with echoes from the E_s layer refer to conditions at that level. Results obtained with echoes from the *F* layer probably represent velocities in the *E* layer just after it is penetrated, but at other times of the night they may refer to the *F* layer itself.

For the purpose of a preliminary discussion all the results are grouped together. It is, however, of great importance to obtain more information about the heights at which the measured movements occur.

(ii) *Relation of the observed velocities to movements of the air*—The movements which are observed are primarily those of inhomogeneities of electron density. There are three reasons why such motion may not correspond to motion of the air. Firstly the electrons might not follow the motion of neighbouring air molecules. Secondly, supposing electrons do, to some extent, move with the air, it is possible for a wave or wave-like motion in the air to produce travelling disturbances of electron distribution. Thirdly, we may have very localized causes of ionization which move horizontally. We must briefly consider these possibilities in turn.

It is known that in the presence of a magnetic field charged particles can move freely along the field but may have their net movement across the field much reduced. This effect requires their frequency of collision to be less than their angular frequency of free gyration in the magnetic field. But since the N–S and E–W components of velocity were found on the average to be of the same order of magnitude we conclude that movement across the magnetic field in an E–W direction is (apart from the possibilities to be considered) an observed fact. Therefore it does not seem profitable to bring in the effect of the magnetic field at this stage. It may be added, however, that an observed movement of patches of ionization across the magnetic field would not necessarily lead to theoretical difficulties because it is only under special conditions that the magnetic field is able to prevent the movement. Normally the effect is compensated by the appearance of electric fields due to polarization charges.

We now turn to the problem of deciding whether wave-like variations passing through the electron distribution could produce the observed effects. The present observations alone cannot settle this point. But it should be noted that methods which apply to isolated "clouds" of ionization or luminous material, such as radio and visual observations of meteor trails, all give velocities similar both in order of magnitude and in degree of variability to those obtained by the present method.

VOL. 2 1952

The third possibility of moving causes of ionization has to be considered in all experiments which observe ionospheric irregularities or clouds whose origin is not clearly understood. But we have just mentioned that velocities derived from observations of meteor trails are in general agreement with the velocities given by the other methods. Since meteor trails are produced in a fraction of a second there can be no confusion here between the velocity of the cause and the subsequent drift of the trail with the air. It is reasonable therefore to suppose that the present results also refer to real wind-like movements of the air.

(iii) *Measurements of velocity by other methods*—The non-radio methods by which the wind at heights near 100 km has been deduced include observations of meteor trails (OLIVIER [12]), polar night clouds (STÖRMER [13]) and self-luminous clouds (HOFFMEISTER [14]). While the first two phenomena occur mainly at heights between 80 and 90 km, the observations of HOFFMEISTER apply to heights of about 120 km well into Region *E*. These methods give velocities of the order of 70 m/sec. Their results however apply to night-hours only. OLIVIER found a predominance of movements to the west at night after midnight over W Europe in agreement with Fig. 4. HOFFMEISTER gave the most frequent directions as towards the N or NE in summer, while in winter they were towards the S as well as towards directions lying between N and E.

Another radio method which has given results applying to heights in or close to the *E* Region of the ionosphere is that of MANNING VILLARD and PETERSON [15], who have used radio echoes from meteor trails. They obtained velocities equal to, or slightly smaller than, those given here. There are some differences in the observed directions, but their published observations are not sufficiently extensive to warrant a detailed comparison of results. It is significant that they found some evidence that the velocity is different at closely adjacent heights. HOFFMEISTERs results for the winter months led him to put forward this same suggestion.

(iv) *Regular changes in velocity*—We turn now to a consideration of the smoothed variation throughout the 24 hrs which appeared to be repeatable on consecutive days. The results of harmonic analysis given in the previous section (Fig. 7) showed that a variation with a period of 12 hr was present in both the N–S and E–W components. It is interesting to ask how this harmonic is related to possible tidal movements in the atmosphere. One component of motion having a significant average 12 hr variation might not be connected with a genuine tidal movement. But the mean vectors in Fig. 7 a and b show that roughly equal harmonics are present in the N–S and E–W components, and that these are about 3 hours out of phase in the sense that corresponds to a clockwise rotation of the horizontal velocity vector. This is what would be expected for a tidal motion in the northern hemisphere. It is difficult to see how such a result could arise spuriously, say from some situation in which the velocity at different heights is measured at different times.

While the magnitude (about 16 m/sec) of the observed tidal motion is not unreasonable from a theoretical point of view, the phase of the observed tide does not fit in well with other evidence. Observations of barometric pressure at the ground give 12 hr variations with a maximum at 1000 hrs corresponding to a horizontal motion at ground level going to the north at 0100 hrs (CHAPMAN [16]). Tidal theory suggests (WEEKES and WILKES [17]) that the tide at about 110 km would be 180° out of phase with this, northward movement occuring at 0700 hrs. The quiet-day solar magnetic variations are explicable in terms of a horizontal tidal motion in the part of the ionosphere where the main electric currents flow, provided its velocity is directed towards the north at 1000 hrs.

The results given here, showing that the semi-diurnal component at the level 100–120 km is towards the north at 0300 hrs local time, do not agree with either of the above deductions. We must presumably conclude that:

(a) the tidal motion near 110 km is not 180 degrees out of phase with that at the ground. The theoretical results may however require re-examination, as WILKES and WEEKES have pointed out that this phase change is sensitive to the assumed distribution of temperature with height.

(b) the solar quiet-day variation of the geomagnetic field is not explicable in terms of a straightforward dynamo theory in which the currents flow in the level 100–120 km.

In seeking to explain movements of the kind here described we must of course remember, as MARTYN has emphasized, that electrodynamical forces on the ionosphere are at least as great as the tidal ones. The theory of such electrodynamical forces is not yet developed far enough to be used in explaining these results.

A similar analysis was carried out to detect the presence of a lunar semi-diumal variation. The results were not significant and the only conclusion that can be made is that the amplitude of such a variation does not exceed 10 m/sec*.

(v) *Possible rapid changes in velocity*—Reference to Fig. 3 shows that the direction of the measured velocity can change appreciably in about half an hour, and may

* Since the paper was written a further analysis, confined to results for *E* region echoes in the daytime, has indicated a lunar tide having a similar phase to that of the solar tide, *i.e.* northward movement occurs about 3 hrs after transit. Approximate amplitudes of 15 m/sec for the lunar tide and 25–30 m/sec for the solar tide are now suggested. J. H. CHAPMAN in Canada first showed that a lunar tide was detectable by this method; the results for Cambridge differ slightly but not significantly from his.

even reverse in this time. On some occasions when long records are studied in detail it is found that successive rapid changes of direction can occur in times of the order of 10 min. Since it would be difficult to contemplate a rapid change of this kind in the mass motion of the air it is desirable to examine the evidence more fully.

We first note that all our observations depend on measurements made on the radio diffraction pattern at the ground, and we have assumed that movements of this are due solely to *horizontal* movements in the ionosphere. But it is possible for a horizontal movement of the pattern to be caused temporarily by other ionospheric changes. One possibility is a tilting movement of the reflecting layer. Another is that there are two close reflection levels which are slightly inclined to each other so as to produce "interference fringes" on the ground. If the levels moved further apart or closer together without changing their relative inclination, the fringes would move over the ground in one direction or the other. This would produce sinusoidal fading at fixed receiving aerials, and an apparent horizontal progression corresponding to the velocity of the fringes over the ground. A situation of this kind could arise if each reflection level corresponded to the level of reflection of one of the magneto-ionic components. For this reason results from records with quasi-sinusoidal fading have been suspected. It should be noted that on the frequency of 2·4 Mc/sec chosen for the experiment only the ordinary component is reflected with appreciable strength by day. Aerials sensitive to only one magneto-ionic component are now being used to overcome this difficulty on other occasions.

VOL. 2 1952

However, when all the doubtful records are rejected, there remain many occasions when the wind-like motion in the ionosphere really appears to change direction. Some of these occasions correspond to the major changes which are shown in the mean variation and are repeatable from day to day. Others correspond to more rapid and less regular changes. Both kinds can occur near midday and do not appear to be accompanied by a change in equivalent height of reflection. Other examples, however, may accompany a change in height of reflection, sometimes within Region E. Reversals sometimes occur, for example, when the echo is returned first from the normal E layer and then from the E_s layer. In order to investigate this matter more fully an experiment is at present in progress in which continuous records of wind velocity and of the equivalent height of reflection are made simultaneously.

It seems likely that reversals of the kind just mentioned are due to the fact that the air is moving in different directions at different levels in the ionosphere as already suggested by other workers. Under these conditions a reversal without a change in the height of reflection might possibly be explained by the sudden production of ionisation at some new level below the reflection level, dense enough to affect the fading so that movement at this new level is recorded.

6. Conclusions

The method of observation here described leads to the conclusion that there are steady horizontal motions in the ionosphere which show a considerable degree of regularity both from year to year and when results at Cambridge, England are compared with those for Washington, D.C., U.S.A. and Montreal, Canada. There is some doubt about the levels at which the motions occur, but circumstantial evidence suggests that the results are applicable to heights in the region 100–120 km.

It is probable that the motions represent real movements of the air mass, but there may be different motions at closely separated heights in *E* region. The order of the velocities, 70 m/sec, agrees with information obtained by other methods. There is a regular semi-diurnal solar component (tide) of amplitude 16 m/sec in the motion. The phase of this solar tide conflicts with a simple dynamo theory of the quiet day geomagnetic variation on the basis of electric currents in the 100–120km region.

Acknowledgements—The Author is indebted to Dr. K. WEEKES and Mr. B. H. BRIGGS for discussions which have been most helpful. The help of Mr. BRIGGS in the experimental observations is also gratefully acknowledged. Professor R. O. REDMAN has kindly allowed the erection of aerials in various parts of the grounds of the *Solar Physics Observatory*, Cambridge.

I wish to thank Messrs. C. D. SALZBERG and R. GREENSTONE of the *Bureau of Standards* and also Mr. J. H. CHAPMAN of *McGill University*, who have been kind enough to send details of their results in advance of publication.

This work was done in the *Radio Section of the Cavendish Laboratory* as part of a programme of radio research supported by a grant from the *Department of Scientific and Industrial Research*. I am also indebted to the Department for a maintenance grant.

REFERENCES

[1] PAWSEY, J. L.; Proc. Camb. Phil. Soc. 1935 **31** 125. [2] MITRA, S. N.: Proc. Inst. Elect. Engrs 1949 (Pt. III) **96** 441. [3] KRAUTKRÄMER, J.; Archiv der Elektrischen Übertragung 1950 **4** 133. [4] BRIGGS, B. H., PHILLIPS, G. J. and SHINN, D. H.; Proc. Phys. Soc. 1950B **63** 106. [5] RICE, S. O.; Bell Syst. Tech. J. 1945 **24** 46. [6] FÜRTH, R. and MACDONALD, D. K. C.; Proc. Phys. Soc. 1947 **59** 338. [7] SALZBERG, C. D. and GREENSTONE, R.; J. Geophys. Res. 1951 **56** 521. [8] CHAPMAN, J. H.; Private communication. [9] FINDLAY, J. W.; in preparation. [10] BRIGGS, B. H.; Proc. Phys. Soc. 1951B **64** 255. [11] HEWISH, A.; Proc. Roy. Soc. (A) 1951 **209** 81. [12] OLIVIER, C. P.; Proc. Amer. Phil. Soc. 1942 **85** 93, 1947 **91** 315. [13] STÖRMER, C.; Astrophys. Norv. 1935 **1** (No. 3). [14] HOFFMEISTER, C.; Z. Meteor. 1946 **1** 33. [15] MANNING, L. A., VILLARD, O. G. and PETERSON, A. M.; Proc. Inst. Radio Engrs 1950 **38** 877. [16] CHAPMAN, S.; Nature, Lond. 1947 **159** 357. [17] WEEKES, K. and WILKES, M. V.; Proc. Roy. Soc. 1948A **192** 80.

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 155 to 159. Pergamon Press Ltd., London

Negative electric fields in mist and fog

J. Alan Chalmers, M.A., Ph.D.

Senior Lecturer in Physics, Durham Colleges, University of Durham

(Received 24 October 1951)

ABSTRACT

The investigation, with transportable apparatus, of the occurrence of negative electric fields during periods of mist and fog has shown that the effect must be ascribed to the emission of an excess of negative ions from overhead high tension power cables in damp conditions, probably during discharges over the insulators.

1. Introduction

Negative electric fields have been reported at Durham during conditions of mist [1] and "high-hanging fog" [2]; the latter condition usually occurs when there is fog at the coast and an E or NE wind.

The phenomenon warrants attention because it appeared to indicate a separation of charge under conditions so different from those of thunderstorms that a totally different mechanism would seem to be required, a mechanism moreover in which precipitation could play no part. The results obtained in the present investigation show that this conclusion is in fact true in an unexpected fashion.

The occurrence of negative fields at Durham in October 1951 prompted investigation with a transportable apparatus for field measurement, which had been constructed for a different purpose, and led to the surprising results to be described.

2. Apparatus

The apparatus used was very simple. A square metal plate of side about 11 cm was fixed on top of a block of paraffin wax, itself fixed to a horizontal metal plate, and the square plate was connected by coaxial cable to the amplifier. By means of strings, a second metal plate somewhat larger than the first can be pulled to cover (without touching) or uncover the lower plate.

The "bound" charge which is on the lower plate when it is exposed to the field must pass through the amplifier when the covering plate is moved, and the amount of this charge gives a measure of the sign and magnitude of the field. The whole apparatus was mounted on the top of a car and could be transported; the strings were pulled from inside.

The amplifier used is the one previously described [3] and has been mounted on a board together with the 2 v LT and 120 v HT batteries; a GB battery is used for calibration. The galvanometer used is a pointer type, shunted to give critical damping. It has been noticed and is worthy of record that it is essential that the amplifier valve should be shielded from daylight, otherwise the sensitivity is much decreased, probably due to photoelectric emission in the valve. The amplifier board was placed inside the car so that the galvanometer deflection could be read when the strings were pulled.

The actual field strength corresponding to a given deflection of the galvanometer depends on the sensitivity of the amplifier and on the "exposure" of the plate, by which is meant the ratio of the number of lines of force ending on the plate to the

VOL. 2 1952

number that would end on a given area of flat ground under the same conditions. The sensitivity can be determined by the calibrating switch, and was found to be fairly constant except if the air was very moist, when the sensitivity was much decreased, probably due to poor insulation of the plate and consequent decrease of the time constant of the grid circuit. The exposure of the plate clearly depends on the exact situation of the car and is governed by the presence of trees, posts, telegraph wires etc., so that it is not possible to gauge accurately what field is represented by a particular deflection. In normal working, 1 division on the galvanometer represents about 10 v/m; when fields were greater than represented by the full-scale deflection of 30 divisions, the sensitivity could be reduced by a decrease in the capacity in the galvanometer line.

It was not found necessary to use any "earth" for the apparatus, the cover and supports and the tin box shielding the amplifier being connected to the negative battery terminals. No effects were ever observed that could be ascribed to charging up of this "earth," and readings under steady conditions were reproducible even after travelling a distance and returning. Comparisions at the Science Laboratories between the movable apparatus and fixed methods of measuring field were always consistent.

3. Measurements

The measurements under consideration were carried out during the first half of October 1951, when anti-cyclonic conditions prevailed, giving usually fog or mist in the mornings, often clearing later. In no case was there any rain or threat of rain during the measurements.

In all 22 journeys of from 4 to 20 miles were made, with from 25 to 115 observations of field on each journey. Most of the journeys were made during the foggy or misty mornings, bur a few were made later in the day, when the weather was fine and in one case a journey was made during a fine evening. In some cases to and for journeys were made over a section of road.

When the weather conditions were steady the same or similar results were obtained at the same place after journeys of up to half-an-hour or sometimes longer, with intermediate results at other places very different. This shows that the results are not concerned with the motion of the car, but were characteristics of the place where measurements were made, at the particular time. But on the other hand, when weather conditions altered, the results at a particular place were very different, showing that the results were not merely some property of the place, but depended also on meteorological factors.

In order to quote figures, the results have been worked out in volts/metre, assuming the exposure to be always the same and using an approximate calibration.

4. Results

The measurements showed negative fields over a greater or less area and of a greater or less magnitude in 20 out of the 22 journeys, and even in the other two cases there were indications of some of the positive fields being below their normal value. Negative fields of over 800 v/m were found on occasion, but only once was the positive field over 250 v/m this being on a railway bridge when a train had just passed beneath. (The increase of positive field due to a train is a well-known phenomenon *e.g.* Israel [4].)

The two occasions on which no negative fields were found and five when only small areas of negative field existed were all during fine weather, when the fog, if any, had cleared. This indicates that the negative fields must be in some way associated with the fog. But is has been found that the negative fields reappear before fog or mist returns, for example negative fields of up to 500 v/m, were found during a fine evening after a morning with fog and intense negative fields and a fine afternoon with no negative field at all. As there was heavy dew in the evening, it is clear that it is the dampness that is to be associated with the negative field.

The actual area over which negative field occurred was found to depend on the wind direction; for example, negative fields at the base of operations, Durham Science Laboratories (A on map) occured for only 8 out of 35 measurements, and with most of these the wind was easterly; on the other hand at Byers Garth Cross Roads (M on map) the field was negative on 22 occasions out of 27, the positive fields occuring only in fine weather.

When winds were westerly or south westerly, it was found that the boundary of negative field was exactly at the place, between H and J on map, where high tension cables, 2 of 66000 v and 1 of 132000 v, crossed the road, the negative field being found to the east; for easterly winds, while the field was negative on both sides of the cables, it was much stronger at H than at J. That this was not a peculiarity of the particular place was shown by finding similar effects at 7 other places where the cables crossed roads in suitable directions, 3 of these being at Q to R, T to U and V to W on the map, the directions of the roads and cables being different in the different cases; in every case where a negative field has been observed, it is possible to ascribe it to high tension power cables to the windward. Table 1 which gives observations for the same places at 3 different times of the same day, in conjunction with the map, shows how the effect depends on the wind direction and on the humidity.

VOL.
2
1952

It is found that quite large effects extend to at least 3 km from the cables down wind; on one occasion with a SW wind when the field was −800 v/m at M, it was found not long afterwards to be between −200 v/m and −300 v/m at points 3 or 4 km to the north-east and over 3 km from the nearest cables.

There is some evidence, though it is not quite conclusive, that the negative fields tend to keep to lower ground, particularly tending to follow a valley in the direction of the wind.

It may be pointed out that peculiar effects are found very close to the cables, probably because the cables themselves provide an alternating field which the amplifier partially rectifies. For this reason, no account is taken of any observations within 50 m of cables.

During foggy weather, when passing close to the pylons supporting the cables, there is often audible hissing due to breakdown at the insulators, and this has been noticed on some of the occasions when the negative fields have been largest.

On a few occasions, particularly after the fog has cleared, results at points fairly close to the cables have shown great variability over periods of minutes; occasionally the field has varied by as much as 100 v/m inside a minute.

It may be remarked that it has seemed that, in fine weather the effects, which are then small, come more readily from the 66 000 lines then from the 132 000 line;

but Table 1 shows that, in the evening it is the 132 000 line which shows the effect most readily though its direction happens to be more favourable in regard to the wind at that time.

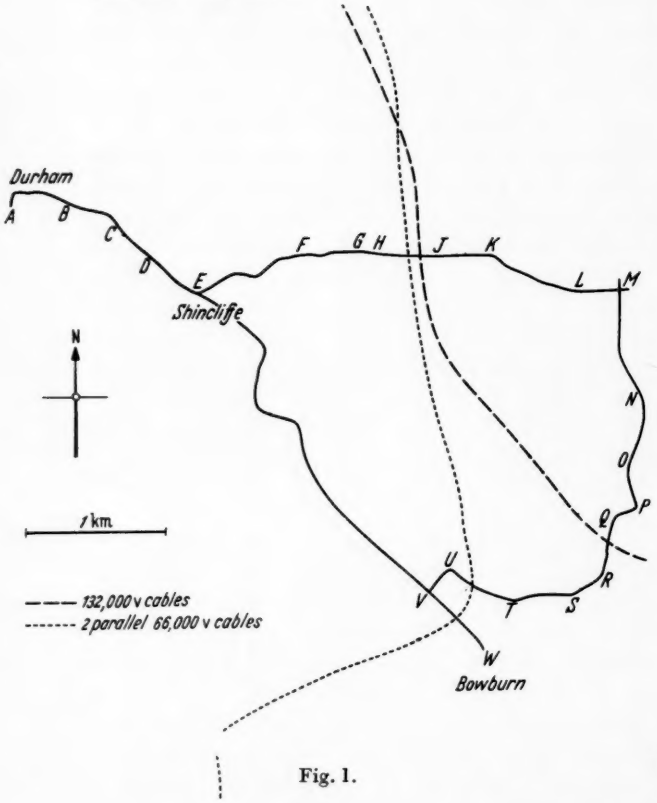


Fig. 1.

Table 1. Fields in v/m at different times and places. October 9th

	10·00 to 12·00	*14·00 to 16·00*	*19·00 to 20·30*
A	+80	+120	+90
B	+80	+60	+50
C	+90	+60	+60
D	+210	+15	+90
E	+130	+10	+90
F	−80	+50	+50
G	−240	+70	+60
H	−180	+60	+120
J	−500	+20	−400
K	−600	+60	−500
L	−180	+80	−170
M	−370	+120	−110
N	−500	+120	−100
O	−900	+100	−220
P	−1000	+100	−120
Q	−730	+100	−260
R	−100	+110	+90
S	−150	+110	+140
T	−250	+110	+160
U	−120	+90	+30
V	−220	+60	0
W	+40	+120	+60
Wind	SE	S	SSW

5. Discussion

The results obtained leave no room for doubt that in humid conditions the high-tension power cables produce effects which, with the help of the wind, are able to cause negative fields. The most probable way in which this could occur would be by the production of negative ions which could be carried by the wind, and it seems that this idea is sufficient to explain all the phenomena described.

Since we know that insulation breakdown does occur at the insulators, it seems reasonable to suggest that this is where the ions originate, but the present observations are not able to confirm this. Nor can they determine whether there are produced ions of both signs, with an excess of negative, or only negative ions.

The most remarkable feature of the results is the distance to which the effect can spread. Effects of discharges might well be expected close to the cables, but it is surprising that they are still large at distances of several km. That this effect has never previously been observed can be ascribed to the fact that transportable field measuring apparatus has hitherto only been used in aircraft. Observations at Kew from 1933 to 1938 [5] may be another example of the same effect.

In order to estimate the magnitude of the effect at pylons, approximate calculations have been carried out which show that the fields observed could be accounted for by currents of the order of microamperes of negative excess liberated at each pylon.

The effects described would be produced by negative ions at whatever height above the earth's surface they might be; if it is true that the negative fields tend to follow valleys, it would seem that the negative ions will be in the lowest layers of the air.

6. Further Considerations

From the present results, it is clearly no longer valid to assume that negative fields necessarily imply a process of charge separations in clouds; so the evidence which suggested charge separations in mist [1] or in low stratus cloud [2] can no longer be accepted, and there appears no reason to believe that charge separation occurs in the absence of precipitation.

Measurements of negative electric fields must be regarded with suspicion until it is established that these fields do not originate from charges produced at pylons. Although negative fields strong enough to cause point discharge have not been found at Durham during the present observations, it is possible that under suitable conditions this might occur, in which case the electrical condition of the atmosphere would be very different from that envisaged for point discharge due to fields originating in the cloud. If, for example, in the measurements of Chalmers and Little [6] the negative field were produced in the same way as in the present observations, then the point discharge current would not give rise to a positive space charge below the cloud, for the positive ions liberated would combine with the negative ions from the pylons. Hence the picture of a field increasing with height would not hold. In other observations, too, the effect of negative ions from pylons would be profound. In any future observations, it will be necessary to test whether pylons are responsible.

If a pylon is emitting negative ions, the corresponding positive charge must go to the earth through the pylon or the cable system. It is probable that a considerable part of the negative charge reaches the earth by conduction currents, but some negative charge will travel upwards in the positive field above the volume of negative ions. The resultant effect is to transfer to the earth an excess of positive charge, exactly as in fine weather, and the consequence of the negative fields discussed here is not, as might be expected at first sight, to help maintain the fine-weather field, but just the reverse.

Further investigation of the effect could only be made by using many continuous recording field measuring instruments situated at various points in relation to the pylons, and also by measuring ion numbers or conductivity with transportable apparatus. It is hoped to investigate the effect during rain with the present apparatus.

The author wishes to express his thanks to the Council of the Durham Colleges for the grant of a sabbattical year which has allowed the present work to be carried out without interruption from other duties.

References

[1] Chalmers, J. A. and Little, E. W. R.; Terr. Mag. and Atm. Elec. 1947 **52** 239. [2] Chalmers, J. A. and Hutchinson, W. C. A.; Nature, Lond. 1949 **164** 68. [3] Chalmers, J. A.; J. Sci. Inst. 1949 **26** 300. [4] Israel, H.; Arch. für Met. Geoph. u. Biokl. 1950 **3A** 1. [5] British National Report on Terr. Mag. and Atm. Elec. (1936–39) p. 13. [6] Chalmers, J. A. and Little, E. W. R.; Terr. Mag. and Atm. Elec. 1940 **45** 451.

VOL. 2 1952

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 160 to 172. Pergamon Press Ltd., London

World-wide simultaneous magnetic fluctuations and their relation to sudden commencements

W. JACKSON

Magnetic Department, Royal Greenwich Observatory, Herstmonceux

(Communicated by the Astronomer Royal)

(Received 22 September 1951)

ABSTRACT

The magnetograms of Cheltenham, Tucson, San Juan, Honolulu and Sitka have been published from 1 January 1946, and distributed by courtesy of the American Coast and Geodetic Survey. Use has been made of 2½ years of these publications to compare the properties of sudden commencements (S.C.s) at Abinger with those of the other stations, also to find out whether there existed other world-wide movements and, if so whether they behaved in the same way as S.C.s. Instances of these movements have been found, and are described below.

The geomagnetic latitutes of the stations are given for reference, and the local times used in the recording of the magnetograms.

	Geomagnetic latitude	*Time behind Greenwich (hours)*
Sitka	+ 60·0	9
Abinger	+ 54·0	0
Cheltenham	+ 50·1	5
Tucson	+ 40·4	7
San Juan	+ 29·9	4
Honolulu	+ 21·1	11

INTRODUCTION

A highly characteristic movement of the elements of the earth's magnetic field which often precedes a magnetic storm or disturbance has received widespread recognition and has been called a sudden commencement (S.C.). The movement has been interpreted as indicating the first impact of a neutral stream of charged particles from the sun on the outer zone of the earth's magnetic field which results, according to the CHAPMAN-FERRARO theory [1], in the setting up of a ring current system outside the earth's atmosphere. It has been used to mark the beginning of certain types of storm—the S.C. storm—and similar impulses which have occurred during the storm have been considered to represent the arrival of further particle streams from the sun.

NEWTON [2] has classified these movements and has shown that, in addition to a minimum near 8 h in the hourly frequency of S.C.s recorded at Abinger, there is a tendency for the movement to be reversed at this time. The S.C. may then be regarded as distorted by a local phenomenon which is, presumably, in or below the earth's atmosphere.

In this paper we will examine a class of world-wide movements of the magnetic elements and compare their properties with those of S.C.s at similar places and times.

They have been selected using the criteria that they are simultaneous at all stations at which they can be observed, within the limits of timing imposed by the available slow-run records, say to a minute, and that the signs of the movements in the three elements are associated in the same way as is found for S.C.s. For convenience of discussion we shall denote this class of fluctuation, when the wave form is similar to but not always identical with a normal S.C., by S.C.-type movement.

The conventional classification will be adopted; an S.C. in which one or more of the elements has a preliminary stroke in the opposite direction to the main stroke is called an S.C.*, a reversed or inverted S.C. one with the strokes reversed from the normal. Representing the changes in horizontal force, vertical force and declination by ΔH, ΔV and ΔD, the normal S.C. for Abinger may be defined as one which occurs at a time of day when anomalous effects are absent and in which ΔH and ΔV are positive, ΔD West.

Since the wave form of S.C.-type movements is sometimes oscillatory, it is appropriate at this stage to differentiate them from a class of oscillation which is easily identified on the Abinger records and in which there is an appreciable phase delay, generally of a few minutes, depending largely on the period of the oscillation, between corresponding movements in the H and D traces. In general these oscillations vary more slowly than the S.C. and show a complex morphology. SCHMIDT has interpreted these movements as being generated by wandering current vortices, and a possible explanation has been given by Price who assumes that they are due to current vortices set up in a non-uniform ionosphere by the induction effects of a changing magnetic field. The perturbations due to these vortices are local and controlled by conditions in or below the earth's atmosphere.

VOL. 2 1952

In the main part of this paper we shall show that many of the characteristic properties of the S.C. perturbation are also found with the S.C.-type movements; this suggests that these properties are characteristic of perturbing field systems generated outside the earth's atmosphere. We shall first discuss some results obtained from a detailed analysis of S.C. movements, then compare similar results obtained from a study of S.C.-type movements. Finally the relation between the two classes of phenomena will be considered.

Section I

CHARACTERISTICS OF S.C. MOVEMENTS

The normal S.C. at Abinger

On the slow-run records the normal S.C. usually appears as a single stroke in each element, the time taken for completion being about 3 min on the average. Sometimes the stroke is complex, suggesting successive increases in H, or oscillations superposed on the main swing, and the pattern is not always a very regular one. Tracings from both slow and quick-run recordings of the same S.C.'s are shown in Fig. 1; it will be seen that the D changes are sometimes independent of H and may be out of phase. V, however, follows H closely in pattern, and the amplitude of the V change bears a very constant ratio (about $\frac{1}{4}$) to that of the H change.

A typical average rate of change of H is 30γ/min, which is frequently exceeded during disturbance so that S.C.s can be regarded as only moderately quick move-

ments. On many occasions the changes in the elements occur very gradually and it is impossible to give any definite time as the commencement of the stroke, while on others an abrupt change may easily be timed to a fraction of a minute.

The "8 h" S.C. at Abinger

If, as is generally assumed, an S.C. is a world-wide effect, any observed diurnal frequency at a given station may be attributed to a suppression or modification of the movement at certain local times of day. The structure of the S.C. has therefore been examined on the Abinger magnetograms for possible modification at particular times. It has been found that there is a period, roughly from 6 h to 12 h local time, during which the structure is anomalous, most pronouncedly so near 8 h; no other period shows such anomalies.

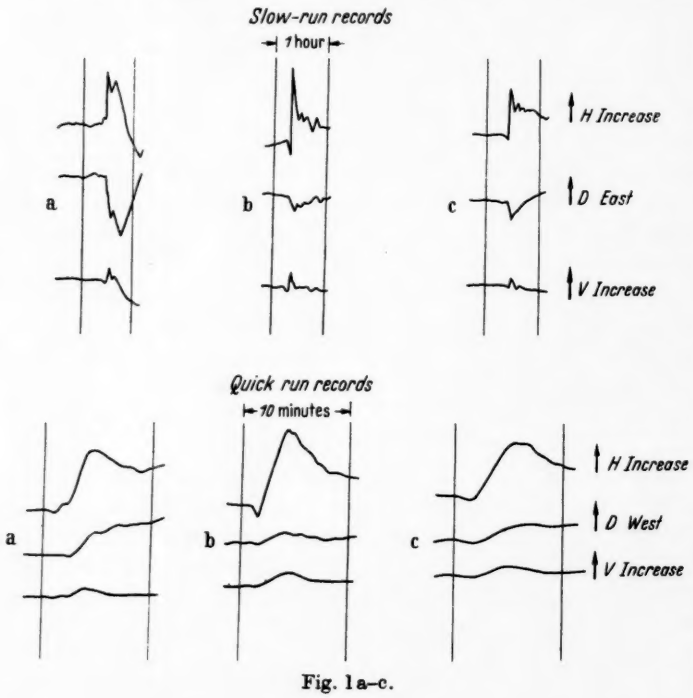


Fig. 1a–c.

During this period ΔD is almost always reversed in direction from the normal, the resultant change being East. The effects in H, however, are variable; sometimes the H trace assumes the appearance of a reversed S.C.*, at other times of a single positive stroke or the beginning of a set of oscillations. Fig. 2, taken from slow-run recordings, gives examples of how a normal S.C. at Tucson or Cheltenham may appear on the Abinger trace if it occurs near 8 h local Abinger time.

More detailed information on the form of the stroke can be obtained from quick-run records, and some S.C.'s which occur near 8 h are reproduced in Figs. 3a to 3e to illustrate the changes which take place. It will be seen that the pattern is a complex one and not always very uniform. For example, the resultant change in D is usually East, but in Fig. 3e it is West, possibly because the S.C. occurs so early. The most interesting features shown by the H trace are the initial movement consisting of a quick rise and fall, completed in about 1 min, the subsequent decrease, not always present, and the presence of oscillations. The initial movement is rather remarkable, not only in that the time for its completion is fairly constant, but that the rate of increase should reach values of 5 γ/sec, equal to the greatest rates of change found at Abinger during storms. While this movement is very characteristic

of Abinger 8 h S.C.'s—sometimes it is the only movement observed—there are occasions when it is absent, as in Fig. 3d. The changes in V appear to be rather closely related to those in H, reduced in magnitude, but there is no such similarity in pattern between the D and H changes.

Comparison with the tracing of a normal S.C. in Fig. 3f shows that the structure of the "8 h" S.C. is radically different, and is not merely a reversed S.C.*. Moreover, the form is not what one might expect from a simple process of ionospheric screening which would have the effect of slowing down the rate of change. In this connection, however, it is possible that the initial quick stroke, as well as the oscillations found associated with these "8 h" S.C.'s, are secondary effects which have their origin in induced vortex currents which are set up in the ionosphere at the time.

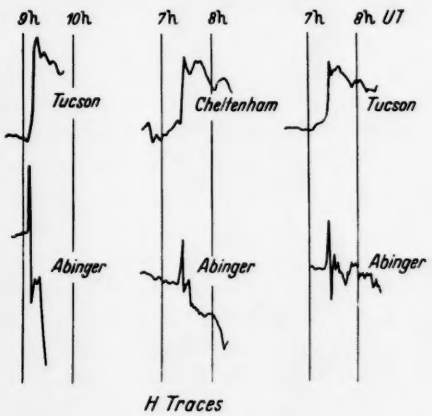


Fig. 2.

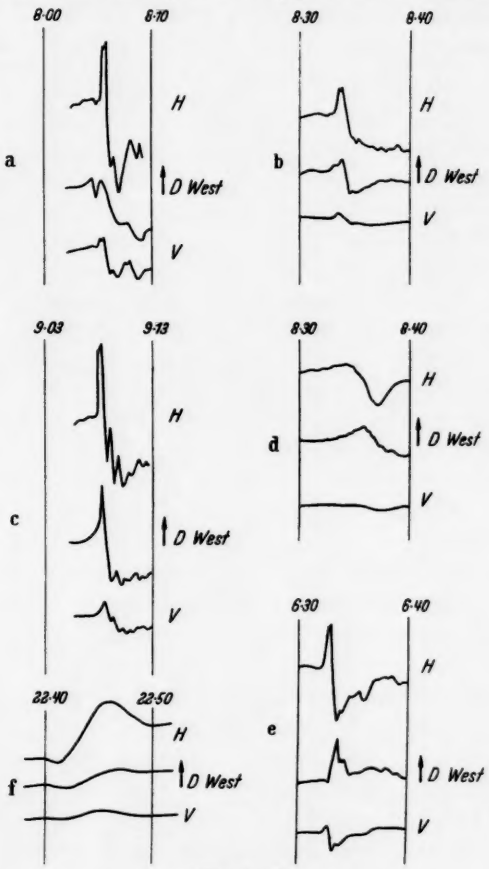


Fig. 3a–f.

Some features of S.C.'s at other stations

Before examining S.C.'s at other stations for possible local effects similar to those at Abinger, it is necessary to establish the changes which take place for the normal S.C. They may be summarised:

	ΔH	ΔD	ΔV
Abinger . .	Positive	West	Positive, about $\frac{1}{4}$ ΔH
Cheltenham	Positive	East	Very small, usually negative
Tucson . .	Positive	West (occasionally East)	Small positive
San Juan .	Positive	Very small	Small negative
Honolulu .	Positive	Very small	Positive, varying from 0·4 to 0·9 of ΔH

VOL. 2 1952

The variable changes in declination at Tucson and the differences in sign between stations supports the conclusion previously drawn [3], that the sign of ΔD is more variable than that of ΔH or ΔV for a given station and may be of opposite sign at distant stations.

For the period near 8 h local time ΔD shows some interesting effects. While at Abinger and Tucson it usually changes from West to East, at Cheltenham it remains East. At San Juan and Honolulu, where ΔD is normally very small, only at this time does it become appreciable in magnitude and its direction is then East. A point of interest is that at Cheltenham ΔD is not reversed near 8 h local time, but the available evidence points to its being reversed near 8 h U.T. or 3 h local time. This is difficult to account for and an analysis of Cheltenham S.C.s over a much longer period would be necessary to confirm this.

In the absence of quick-run recordings, the detailed changes which take place in H at these other stations cannot be studied, but some general results from the slow-runs may be given. At San Juan and Honolulu reversal effects are not found, and at Tucson they are infrequent. There is evidence, however, of a reduction in the magnitude of the S.C. near 8 h. Reversals, on the other hand, do take place at Cheltenham and, as far as can be judged from the slow-run records, they are very similar to those that occur at Abinger.

It may be concluded that the period near local 8 h is one in which abnormal effects do occur at other stations, although the effects are not the same. In so far as the stations examined are representative of latitude effects, it would appear that the complex effects at higher latitudes are not found at lower latitudes, where only a simple reduction in magnitude of the stroke, suggestive of screening effects, is found.

The relative magnitude of S.C.s at various stations

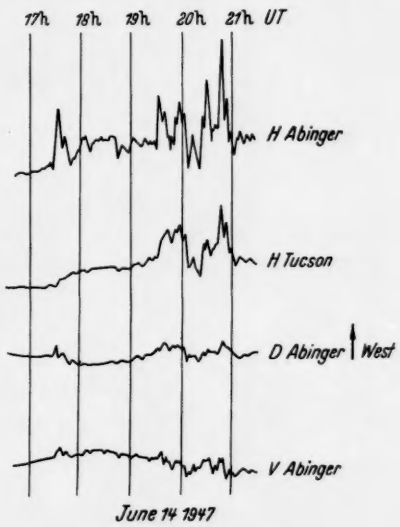
The magnitude of the S.C. is of importance in determing its relation to other types of disturbance, *e.g.* the ring-current type, or in an investigation of screening effects. The changes in H are first considered, since those in D and V are usually small and somewhat anomalous. S.C.s have been selected outside the 6–12 h local time period, *i.e.* free from the suppression or reversal effects which may take place then. This subdivision ignores possible diurnal effects which occur outside this period, but, if they do exist, they are much less marked than the anomalies at this time, and, in any case, the result is accepted only as an approximation. Averaging for all the S.C.s examined, it was possible to arrive at an estimate of the mean relative magnitudes in H of S.C.s recorded at the various stations, the probable errors of the mean being small enough to ensure statistical significance to the result. The average magnitudes of ΔH relative to Abinger as unity were Sitka 1·2, Cheltenham 1·1, Tucson 0·8, Honolulu 0·5, San Juan 0·5.

While the H changes at the different stations are such as would fit a general pattern of decrease towards the equator, no such effect can be detected in the V changes. It has been remarked that the V change at Abinger is about $\frac{1}{4}$ that of H. At Cheltenham, Tucson, San Juan it is very small. At Honolulu it varies greatly, a test sample showing it to range in magnitude from 0·4 to 0·9 that of H, with an average ratio of $\Delta V/\Delta H$ about 0·6, in wide contrast to the remarkably constant ratio at Abinger. The large value of the positive ΔV at Honolulu is surprising; it may

be noted that the diurnal variation in V at Honolulu is also abnormally large. It is uncertain whether there is any connection between the amplitudes of the diurnal variation and that of S.C.'s, though at Huancayo, where the amplitude of the diurnal variation in H is abnormally large, the magnitude of ΔH in S.C.'s is also found to be very large. It is concluded that anomalies exist in the V changes of an S.C., and that the lack of any uniform pattern in the changes suggests that local effects predominate.

The form of the S.C. at different stations

It may be appropriate at this stage to comment on the form the normal (as apart from the 8h) S.C. assumes at various stations. A study of these forms would suggest that each station has a characteristic type of S.C. At the low latitude stations, for example, the S.C. appears as a simple rise in H, and the H trace remains high for an hour or two. On progressing to higher latitudes, the S.C. assumes a more oscillatory character, and the H trace returns much sooner to normal. This effect is most striking at Sitka, where the S.C. is sometimes completely oscillatory, without any resultant rise in H. At medium latitudes the S.C. assumes a form intermediate between these two; often the oscillatory character may be deduced from the "tail" at the top of the stroke, even when none is evident at the beginning.



a

Fig. 4a.

Section II

Characteristics of S.C.-Type Movements

Some selected groups of S.C.-type movements are shown in Figs. 4a to 4c. They conform to the criteria that the separate movements are simultaneous in the three elements, are world-wide in their occurrence, and have the relative signs of H, D and V as in S.C.'s. It will be noted that they are prolonged in time over several hours and that while some have the appearance of a succession of S.C. movements, others are of a quasi-oscillatory character. For the latter, a definite direction cannot be given to H or D, but it is possible to associate a given change of H with a corresponding direction in D. It is worth remarking that oscillations at one station may, at another, appear as movements which would be regarded as typical S.C.'s. For example, the movement at 6 h 50 m at Cheltenham in Fig. 4b is not dissimilar from the oscillation just after 1 h; yet at Honolulu the movement could reasonably be taken as an S.C. (a large depression in H does in fact follow), and at Abinger it behaves exactly as one might expect an S.C. to do at this time. Comparison of magnitudes at different stations is more difficult than for S.C.'s because of the oscillatory nature of the stroke, but it has been ascertained that the relative change of magnitude is similar to that for S.C.'s.

VOL. 2 1952

It will be noted in Fig. 4a that the Tucson trace is partly suppressed between 17 h and 19 h (10 h and 12 h local time). After this the similarity between the movements at Abinger and Tucson is pronounced. Further examples of these movements

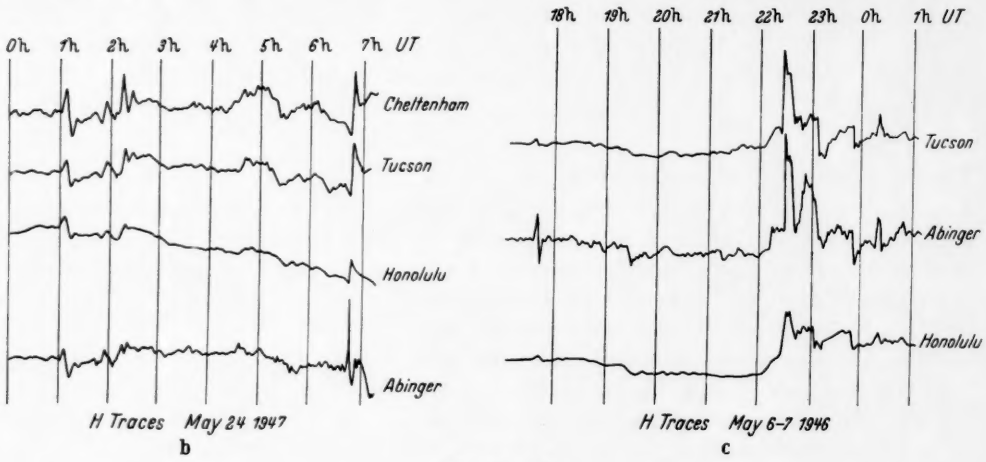


Fig. 4 b and c.

at times near local 8 h are shown in Fig. 5a to 5f and summarised in Table 1. It is important to note that, at the local times when S.C.'s show suppression or reversal effects, S.C.-type movements do so also. In Fig. 5a for example, the Cheltenham *H*

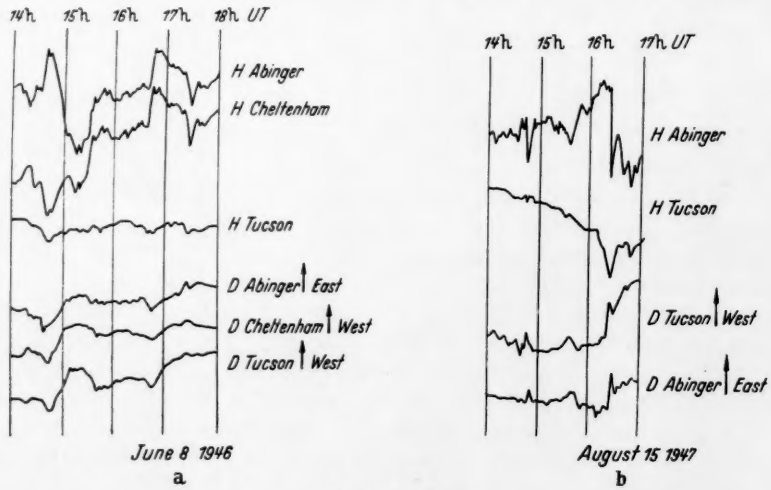
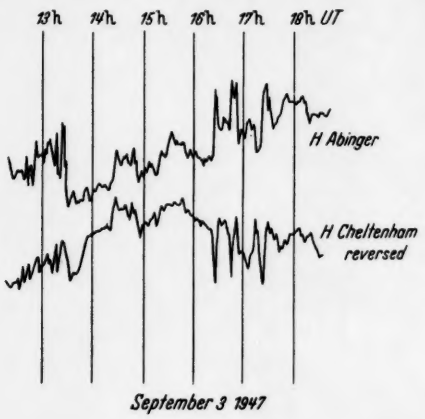


Fig. 5a and b.

trace is reversed from Abinger at local time 9–10 h (U.T. 14–15 h) and normal from 10–13 h (U.T. 15–18 h). The Tucson *H* trace from local time 7–11 h (U.T. 14–18 h) shows both reversal and suppression effects. The signs of the movements on the *D* traces are the same as for S.C.'s at the appropriate local times, both the

Cheltenham and Tucson traces being reversed from Abinger. Reversals of H are not very common near local 8 h at Tucson and the more frequent effect, partial suppression of H with reversal of D, is illustrated for local time 7–10 h (U.T. 14–17 h) in Fig. 5b. A good example of reversal of H at Cheltenham is contained in Fig. 5c; the Cheltenham trace has been inverted



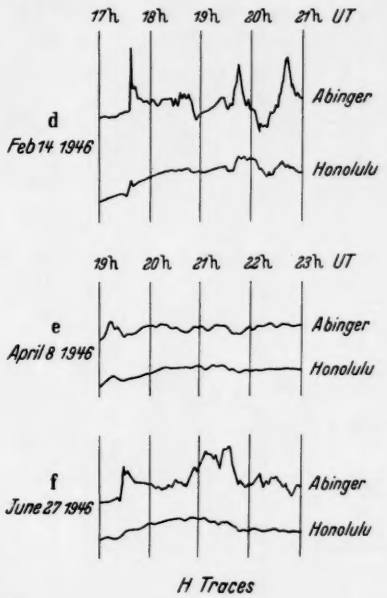


Fig. 5c and d–f.

to show the striking similarity between the Abinger trace and the Cheltenham reversed, from 13–16 h U.T. (8–11 h local Cheltenham time).

At the southern stations (Honolulu and San Juan) reversals of H have not been found, and this holds equally for S.C.-type movements as for S.C.'s; the movements in D are usually so small as to make recognition of changes in any particular direction difficult. Suppression near local 8 h appears to be variable. In Figs. 5a and 5f it is marked but in Fig. 5e it is hardly noticeable, though in each case the period contains the local 8 h.

It has previously been pointed out in Section I that at stations like Abinger and Cheltenham the H change in the 8 h S.C. is variable and that, though pseudoreversals may occur, the stroke may degenerate into a single rapid rise and fall, or the beginning of a set of oscillations. It might be expected that similar effects could be found in S.C.-type movements, and Fig. 6a has been inserted to show that this can be so; the separate movements in H at Cheltenham and Tucson appear as a set of continuous oscillations between 6 h and 8 h at Abinger.

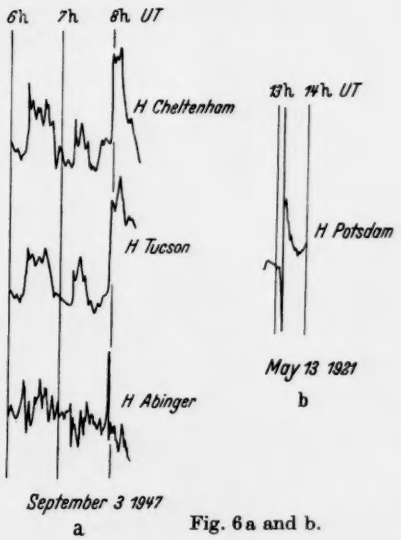


Fig. 6a and b.

VOL. 2 1952

Table 1

Fig.	*Station*	*Date*	*Time (U.T.)*	*Local time at station*	*Horizontal component (compared with Abinger)*	*Declination (compared with Abinger)*
5a	Cheltenham	8 June 1946	14–15 h	9–10 h	Reversed	Reversed
	Cheltenham	8 June 1946	15–18 h	10–13 h	Similar	Reversed
5a	Tucson	8 June 1946	14–15 h	7–8 h	Reversed and partly suppressed	Reversed
5b	Tucson	15 Aug. 1947	14–17 h	7–10 h	Similar in sign but greatly suppressed	Reversed
5c	Cheltenham	3 Sept. 1947	16–18 h	11–13 h	Similar	Very small movement
	Cheltenham	3 Sept. 1947	13–16 h	8–11 h	Reversed	Very small movement
5d	Honolulu	14 Feb. 1946	17–21 h	6–10 h	Partially suppressed	Very small movement
5e	Honolulu	8 April 1946	19–23 h	7–12 h	Similar, with only slight suppression	Very small movement
5f	Honolulu	27 June 1946	17–21 h	6–10 h	Strongly suppressed	Very small movement

The relation between S.C.-type movements and actual disturbance has not been considered, apart from the fact that they may be found during storminess. On occasions, however, small disturbance could be recognised as consisting mainly of these movements. For example, the disturbance lasting from 13 June to 17 June 1947, and beginning with an S.C. at 17 h 50 m on the 13th, was made up in great part of S.C.-type oscillations.

Section III

DISCUSSION

In the CHAPMAN-FERRARO theory of magnetic storms the S.C. is generated in the initial processes involved when the neutral stream from the sun interacts with the earth's magnetic field. PRICE [4] has suggested that the shape of the observed perturbation is controlled by screening effects in the ionosphere. The shape may also be modified by local induced current systems, partly controlled by induction effects.

On CHAPMAN and FERRAROs model, the variation in field will be controlled mainly by the behaviour of the stream as it meets the earth's magnetic field or by the ring-current system. The primary field, according to PRICE, will be modified by screening effects which will be most effective near the sub-solar point. In both cases the field vector will be parallel to the earth's magnetic axis, producing maximum ΔH, zero ΔV, at the equator, zero ΔH, maximum ΔV at the poles. The secondary field implicit in PRICEs model will be controlled by a rather complex system of currents in the ionosphere which shows a fairly regular variation in the distortion of the S.C. perturbation. It is useful to examine the data to see whether these phenomena are in fact present.

The first point suggested by the analysis is that the S.C. appears in its most simple form and has its smallest magnitude at lower latitude stations (Huancayo is an exception but this station is anomalous in many ways). The amplitude of ΔH is smallest there, and, while the V and D changes are irregular, there is certainly no evidence of an increasing ΔV away from the equator. In this respect the form and magnitude of the S.C.

does not agree very well with that expected from an equatorial ring-current. In addition, the 8 h effects show that, whatever is the fundamental field which causes the S.C. it is not always the dominating one.

The second point worth remark is that there is no marked difference between the magnitude of night and day S.C.s. From this it may be concluded (1) that the S.C. is not to be attributed to currents flowing in the ionosphere which depend on its conductivity and (2) that screening effects of the ionosphere cannot be greatly important in determining the observed magnitude of the S.C., since the conductivity of the ionosphere at midday may vary by a factor of 50 or so over that at night. In particular the sub-solar point is not the region where maximum screening effects are observed. While there is evidence of suppression at lower latitudes it seems more nearly related to local 8 h than to noon, and at the latitude of Abinger the effects found in the local 6 h to 12 h period cannot be explained by a simple screening process. The fact, too, that the changes from normal to reversal which take place in S.C.-type movements appear to do so quite suddenly does not suggest ionospheric effects which depend on the position of the sun. On the other hand it is difficult to postulate any external field as the cause of the "8 h" effects, since the changes are not uni-directional. For example, the process of reversal could cause an inverted S.C. occuring at 13 h U.T. at Abinger to appear as a normal S.C. at Cheltenham (local time 8 h).

VOL.
2
1952

The relation between S.C.s and S.C.-type movements

It has been shown that, in their terrestrial effects, S.C.-type perturbations have characteristic features which are indistinguishable from those of accepted S.C.s. The most important differences seem to be that S.C.-type movements are prolonged in time and have no immediate relation to storms whereas the normal S.C. is a single stroke and is regarded as closely connected with storms. The fact that S.C.-type movements may be found on a quiet trace with no measurable depression in H implies that they are not to be associated with ring-current effects.

The connection between S.C.s and storms is, however, not a very definite one, if S.C.s are accepted as such only if they impinge on a quiet trace and precede a storm. It is true that S.C.s can begin storms, in the sense that they start the large oscillations of the storm, but there are also occasions in which a quiet period intervenes between the S.C. and the storm. This is consistent with the hypothesis that an S.C. is the first terrestrial effect of the corpuscular stream, if the formation of a ring current, or the first entry of particles into the auroral zones, is delayed; but there is nothing in the records themselves to indicate whether an S.C. is associated with a disturbance or not. On occasions pairs of S.C.s, separated by a short interval of time, have been found preceding disturbance. A further difficulty is involved, in that a quiet trace at one station may be disturbed at another during the same interval of time; thus an S.C. at Sitka is often preceded by disturbance of the polar type, sometimes of such magnitude that the S.C. is almost obscured, when southern stations record a quiet trace. The general conclusion seems to be that the beginning of a storm is not always a very definite point in time. Indeed, it might be taken as one of three distinct phases—the first impact of the particle stream, the formation of a ring-current, or the entry of particles into the earth's atmosphere.

The fact of double and multiple S.C.s raises the question of the actual structure of the stream—whether variations in the speeds of the constituent particles could produce bunching of these particles in the path of the stream. If so, successive variations in the density of the stream as it overtakes the earth might produce terrestrial perturbations of the kind described as S.C.-type.

Inverted S.C.s

Apart from the movements in the local 8 h zone, which, as has been shown, are not true reversed S.C.s, there remains a small percentage of inverted S.C.s *i.e.* movements which are reversed in direction from the normal S.C. The number in NEWTONs analysis would be about 4%. The negative ΔH dissociates them from the first effect of the advancing stream and therefore from true S.C.s, though they may possible be connected with S.C.-type movements. Examination of the Abinger records for the analysis referred to provided some evidence for this since they were seldom, if ever, directly related to the commencement of a storm, and they were sometimes recognised as the terminal movement of a group of S.C.-type. The picture suggested is that the H change in S.C.s is always positive, unless modified at the anomalous local times; other reversed S.C. movements are associated with S.C.-type movements rather than S.C.s.

S.C.s with preliminary impulses

A feature of S.C.*s is that they show a tail not only at the beginning but at the end of the stroke—that they are in fact of an oscillatory nature. A good example is the very large S.C. recorded at Potsdam on 13 May 1921 (Fig. 6b) where the resultant rise in H is small compared with the total swing. Examination of S.C.*s showed that the preliminary impulse was much less pronounced at San Juan and Honolulu than at the other stations, often disappearing entirely. CHREE [5], in his description of Antarctic S.C.s, pointed out the oscillatory nature of the S.C. there, and HARANG [6], in a discussion of selected S.C.s from Tromso, Abisko, Lovo, and Niemegk, all high or medium latitude stations, believed that all S.C.s have the preliminary impulse, not always recognised because of defective recording. If S.C.*s are in fact more frequent at higher latitudes, where the ordinary trace is also subject to superposed oscillations, it may be that the tails are due to oscillations associated with, or superposed on, the rise of H. The diagram of a quick-run recording of a typical S.C.* in Fig. 1c shows that it is not of a symmetrical sine-wave type and is of such a nature that it could be explained as an oscillation associated with a rising H.

The diurnal variation of S.C.s

The analysis carried out showed that an S.C. at one station could nearly always be detected at the other stations. Recognition was sometimes difficult because of the modification in form in the local 8 h zone or because the S.C. occurred during disturbance, generally of the polar type which faded out at the southern stations. This was particularly so for small S.C.s. It is apparent therefore that a diurnal frequency in S.C.s may be ascribed to various causes, including the observer's method of selecting S.C.s and his knowledge of what form the S.C. is likely to take at different local times, the amount of disturbance at the time, and the size of the S.C., since the smaller its magnitude the greater the chance of its being lost in dis-

VOL
2
195

turbance and the more difficult it is to detect when modified in structure. In view of this, interpretation of diurnal frequency curves becomes a difficult matter, and it is not easy to draw conclusions from them. For example, S.C.s selected by the criterion that they impinge on a quiet trace and precede disturbance will show a different diurnal variation from similar strokes which are not restricted in this way and which may occur during disturbance. Again, a minimum near local 18 h such as found by NEWTON [2] may be due either to modification of the S.C. at this time, or to increased general activity. There is no evidence of the first, from the magnetograms, but it is known from the K numbers that enhanced activity does occur in the afternoon hours. That S.C.s may be reported at one station and not another may be illustrated from the Alibag results for 1905–44 [7] in which a number of S.C.s were reported at Alibag and not at Abinger, and vice versa. In fact, all the Alibag S.C.s could be recognised at Abinger but were not reported because they occured in the middle of, or after, polar type disturbance probably very weak at Alibag, or because they were modified near local Abinger 8 h.

SUMMARY

The structure and general characteristics of S.C. movements have been examined at different times of day, on quick and slow-run recordings at Abinger, and on slow-run recordings at other stations. The essential differences which exist between stations, especially at the anomalous 6–12 h period, stresses the desirability of quickrun recordings, not only for a more accurate timing of S.C.s, but for an investigation into the causes of these anomalies.

It has been shown that world-wide perturbations, lasting for several hours and often dissociated from large disturbance, may be found on the magnetograms, and that their properties are the same as, or very similar to, those of S.C.s. Some difficulties have been pointed out in defining an S.C. from the magnetograms alone, and in interpreting diurnal frequency curves.

The advantage of publishing copies or tracings of S.C.s and groups of S.C.-type movements is obvious. In this way observers could deduce the local times when anomalous effects, or suppression, take place.

It is pointed out that the period ($2\frac{1}{2}$ years) over which this investigation was made was very short, and that the conclusions are of a general nature and intended as suggestions that may be verified by co-operating observatories working over a much longer period of time.

Acknowledgements—My thanks are due to the Astronomer Royal for facilities to undertake this work and to Mr. W. R. PIGGOTT, Radio Research Station, Slobgh, for his most valuable interest, criticism, and suggestions.

REFERENCES

[1] CHAPMAN, S. and FERRARO, V. C. A.; Terr. Mag. 1931 **36** 77 and 171; Terr. Mag. 1932 **37** 147. [2] NEWTON, H. W.; Mon. Not. Roy. Astr. Soc., Geophys. Supp. 1948 **5** 159. [3] CHAPMAN, S. and BARTELS, J.; Geomagnetism, p. 298. Oxford, 1940. [4] ASHOUR, A. A. and PRICE, A. T.; Proc. Roy. Soc. A 1948 **195** 198. [5] CHREE, C.; Studies in Terrestrial Magnetism, p. 130. London. [6] HARANG, L.; Geofysiske Publikasjoner **16** (No. 12) 39 (Oslo). [7] CHAKRABARTY, S. K.; Nature, Lond. 1951 **167** 31.

VOL. 2 1952

APPENDIX

Study of the Sitka magnetograms showed that, while the H component in S.C.'s is modified near the local 8 h in a manner similar to that found for Abinger and Cheltenham, the shape of the D trace in many instances bears a distinct resemblance to the H component of the same S.C. as it appears at these stations. Periods of S.C.-type disturbance at Sitka have since been investigated and it has been found that, during a certain period of local Sitka time, the same remarkable similarity between the Sitka D trace and the Abinger H trace holds for these movements also. An example is given in Fig. 7.

The time during which this effect may be observed appears to be between about 8 h and 13 h local Sitka time. Outside this range the relationship breaks down. In

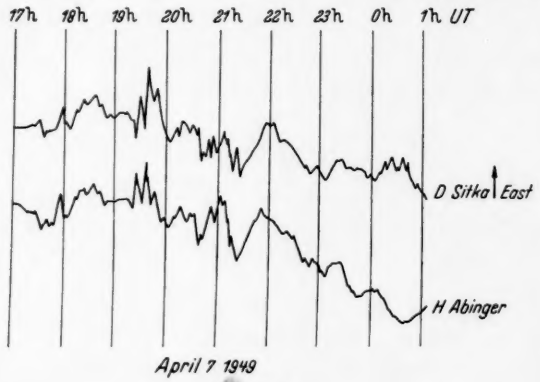


Fig. 7.

this respect the effect is slightly later in time than the "8 h" effects which have been described for the more southern stations. In one or two instances close correspondence was found between both Sitka D and Abinger H, and Sitka H and Abinger D, but in general it is the similarity between Sitka D and Abinger H which is striking. The magnitude of the D movements at Sitka, expressed in γ, is about the same as the corresponding H movements at Abinger. In direction, D East at Sitka corresponds to increase of H at Abinger.

In terms of current flow, the relationship would seem to suggest that an E-W component at Abinger or Cheltenham is diverted to a flow in or out of the auroral zone near Sitka at this particular local time. If this is so, it gives some hint that the local 8 h effects described for other stations may be related to changes which take place in the conductivity of the auroral zone. It has been pointed out that these 8 h effects in S.C.'s and S.C.-type movements are not readily associated with changes in upper atmosphere ionisation due to the sun, or with an external field. If, however, the ionosphere in the auroral zone has abnormal conductivity between 8 h and 12 h local time during this type of disturbance, the effect of a rapidly changing external field would be to produce induced current vortices which could be responsible for the reversal effects observed at stations like Abinger and Cheltenham. It is of interest that the local 8–12 h period in the STÖRMER theory is anomalous in the sense that it is the only part of the zone into which electrons do not enter, but there is no evidence for relating the conductivity with such a cause.

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 173 to 182. Pergamon Press Ltd., London

Über Schwankungen des Ozons in der Troposphäre und Stratosphäre

On Fluctuations of Ozone in the Troposphere and Stratosphere

Erich Regener

Stuttgart und Weißenau*

(Received 20 September 1951)

ABSTRACT

The complexity of the problem of atmospheric ozone is exhibited in the fluctuations of the ozone concentration at all heights. In the layer of air immediately adjoining the ground, the ozone content always sinks to zero when the air stagnates, due to the effect of the ground, which is very destructive to ozone. In the troposphere, advection plays the main part in causing fluctuations. Widely varying vertical ozone distributions have been found recently by means of the spectrographic method in balloon ascents. Te explain this, large-scale horizontal and vertical movements of air at great heights must be assumed.

VOL.
2
1952

Über Schwankungen des Ozons zu sprechen, das bedeutet, über das ganze komplizierte Problem des atmosphärischen Ozons zu sprechen. Wir wissen, daß die Schwankungen zeitlicher und örtlicher Art sind. Sie treten in allen Höhen auf, wo überhaupt Ozon vorhanden ist. Die Amplitude der Schwankungen ist überall verschieden und wenn zeitliche Perioden vorhanden sind, dann überdecken sich tägliche und jahreszeitliche Einflüsse mit örtlichen Effekten. Wir können daher erwarten, daß in den Schwankungen alle Faktoren mehr oder weniger wirksam sind, die für das ganze Ozonproblem wesentlich sind: die Geschwindigkeit der Bildung und der Zerstörung des Ozons, der vertikale und horizontale mechanische Transport ozonhaltiger Luftkörper, ihre örtliche Herkunft und sogar ihre Geschichte. Bei dieser Sachlage erscheint es nützlich, möglichst viel Beobachtungsmaterial zu sammeln, einzelne markante Situationen zusammenzufassen und sie vorsichtig zu interpretieren. Von kleinen zu großen Höhen fortschreitend mögen daher im folgenden die Verhältnisse an älteren und neueren Messungen zusammenfassend diskutiert werden.

Am einfachsten erscheint uns heute der Fall der Schwankungen des Ozongehaltes der Luftschichten in der Nähe des Erdbodens. Hier sind die Schwankungen auch prozentisch am größten; denn hier tritt wenigstens zeitweise aber recht häufig und bei gegebenen Bedingungen auch regelmäßig, der Ozongehalt Null auf. Das ist schon ein Hinweis dafür, daß der Erdboden die Stelle einer starken Ozonzerstörung ist. Das ist nach der chemischen Struktur des Ozons auch nicht anders zu erwarten. Überall auf der Erdoberfläche, natürlich besonders dort, wo Pflanzenwuchs ist, sind oxydable Substanzen vorhanden, an denen sich das Ozon aufzehrt. In gleicher Weise wirken auch Staub, der eventuell in der Luft vorhanden ist und auch die industriellen Gase in unseren Städten. Auch im Kontakt mit vollkommen inaktiven festen Körpern verbraucht sich das Ozon durch katalytisch wirkende Einflüsse. Es gibt ja auch in einem vollkommen sauberen Glasgefäß einen spontanen Zerfall des Ozons, der natürlich mit der Temperatur stark anwächst. Das gleiche gilt auch für die freie und vollkommen reine Atmosphäre.

* Vortrag gehalten auf der Tagung der *JGGU* in Brüssel am 30. 8. 1951.

Wenn also Ozon in der bodennahen Luft gefunden wird, dann muß es von oben kommen, aus den Schichten, in denen sich eine Gleichgewichtskonzentration des Ozons durch das Zusammenwirken von ozonbildenden und ozonzerstörenden Wellenlängen des Sonnenlichtes und von rein thermischer und deswegen auch stark temperaturabhängiger Ozonzersetzung ausbildet. Die Beobachtungen zeigen, daß die mit den genannten Faktoren berechnete theoretische Ozonverteilung in allen Höhen durch Advektion und Konvektion lokal stark verändert wird. In Höhen unterhalb von etwa 20 km ändert sich aber der relative Ozongehalt bestimmter sich bewegender Luftmassen nur sehr langsam, d. h. erst nach einer Reihe von Tagen merkbar [1]. Erst in Berührung mit dem Erdboden oder mit staubhaltiger Luft sinkt der Ozongehalt schnell ab, eine Erscheinung, die natürlich in der reinen Luft über ausgedehnten Wasserflächen ausbleiben muß. Eine photochemische Bildung von Ozon in der Nähe des Erdbodens ist nach unseren heutigen Anschauungen ausgeschlossen. Höchstens bei starken elektrischen Feldern, wie sie bei Gewittern auftreten, ist eine Ozonbildung durch gelegentliche Spitzenentladungen möglich. Dieselben dürften aber wohl so selten auftreten, daß sie für eine Ozonbilanz nicht in Betracht kommen.

Der Transport des Ozons aus den von der zerstörenden Wirkung des Erdbodens unbeeinflußten Luftschichten bis zum Erdboden herab kommt offenbar durch turbulente Luftaustauschvorgänge zustande. Diese Turbulenz kann sehr verschieden stark sein, sie kann ferner großräumig oder kleinräumig sein. Aus dem Zusammenhang aber zwischen der Art und der Stärke der Turbulenz und dem Rückgang der Ozonkonzentration in Bodennähe muß man auf die Stärke der desozonisierenden Wirkung des Bodens selbst schließen können. Die Beobachtungen zeigen, daß nicht nur in vollkommen ruhender, stagnierender Luft, sondern auch bei schwachem, laminar über das Land streichendem Wind der Ozongehalt in der Nähe des Erdbodens in einem ebenen Gelände in kurzen Zeiten, d. h. in 1–3 Stunden auf Null herabgehen kann. Die Beobachtungen zeigen auch, daß die Ozonzerstörung am Erdboden besonders schnell vor sich geht, wenn am Abend und in der Nacht durch eine bodennahe Inversion eine Luftschicht von geringer vertikaler Ausdehnung gegen den Austausch mit höheren Schichten abgeschlossen wird. Wenn in einem solchen Falle morgens die Sonne über dem Horizont erscheint, dann genügt die dadurch hervorgerufene geringe und kleinräumige Thermik, um die Bodeninversion relativ schnell zu zerstören und unbeeinflußte, ozonreichere Luftmassen aus geringer Höhe herabzuschaffen. Die ozonzerstörende Wirkung des Erdbodens reicht in flachem Lande wohl nicht hoch hinauf, schätzungsweise auf 50–100 m natürlich je nach dem Grad der Luftdurchmischung, bei großräumigen Temperaturinversionen und großräumiger Kaltluft in Talkesseln natürlich auch weit höher.

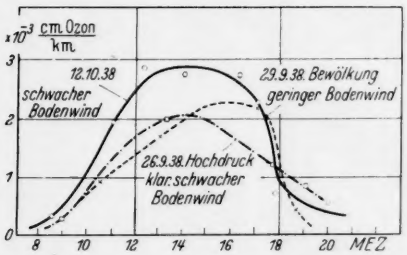
Bei stärkerer Turbulenz, insbesondere bei westlichen Gradientwinden, geht natürlich der turbulente Austausch viel schneller vor sich und reicht auch viel höher hinauf. Die Beobachtungen zeigen, daß dann der Transport unzerstörten Ozons von oben die ozonzerstörende Wirkung des Erdbodens stark überwiegt und der Gang des Ozongehaltes am Boden ziemlich derselbe ist wie in geringer Höhe in freier Luft.

Natürlich ist dieser Vorgang von den lokalen Verhältnissen, von der Orographie des Ortes stark abhängig und läßt sich am besten an Hand von Beispielen erläutern. Ich zeige daher zunächst einige Messungen, die schon kurz vor und auch im Kriege von meinen Mitarbeitern in der Nähe von Friedrichshafen am Bodensee ausgeführt worden sind. Die Verhältnisse sind dort an der Grenze zwischen der großen

Wasserfläche und dem flachen Ufer besonders einfach und die einzelnen wirksamen Faktoren treten denn auch bei den Messungen sehr deutlich in Erscheinung.

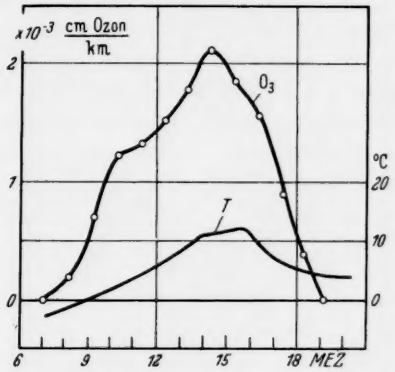
Abb. 1 zeigt den von R. Auer [2] nach einer schnell arbeitenden Jod-Kalium-Methode bestimmten Tagesgang des bodennahen Ozons an schönen, windschwachen Herbsttagen. Man erkennt einen hohen Ozongehalt tagsüber, der bald nach Sonnenuntergang auf einen unmeßbaren Betrag absinkt, um nach Sonnenaufgang bei zunehmender Thermik der bodennahen Luft wieder anzusteigen.

Daß es die von dem erwärmten Erdboden ausgehende thermische Turbulenz der untersten Luftschichten ist, die nach



Tagesgang des bodennahen Ozons in Friedrichshafen/B (nach R. Auer)

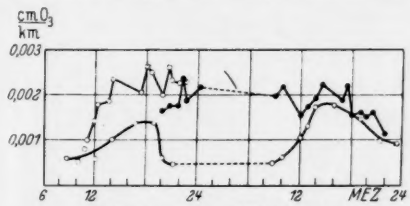
Abb. 1.



Ozon und Lufttemperatur in Friedrichshafen am 12.3.1940 Nach A. Ehmert

Abb. 2.

dem nächtlichen Minimum Ozon aus der Höhe herabbringt, ersieht man auch aus den von A. Ehmert [3] gemessenen Kurven in Abb. 2, die den mit der Thermik parallelen Gang zwischen der Lufttemperatur und dem Ozongehalt mit einem Maximum zwischen 14 und 15 Uhr zeigen. Bemerkenswert ist auf dieser wie auch bei den Kurven in Abb. 1, das steile Abfallen des Ozongehaltes mit sinkender Sonne gegen Abend. Offenbar bildet sich sehr schnell eine Temperaturinversion aus, die dicht über dem Boden liegt und eine dünne Luftschicht gegen den Austausch mit den etwas höheren Schichten abschließt. In dieser dünnen Schicht kann dann die Ozonzerstörung schnell vor sich gehen. Dieser Effekt kann unterstützt werden durch den Umschlag der Luftbewegung, die am Abend ozonarme Luft vom Lande zum See führt, am Morgen umgekehrt. Bei geringer Allgemeinbewegung der Luft hat man es am Morgen mit Luft zu tun, die sich über Nacht über dem See befunden hat und ozonzerstörenden Einflüssen kaum ausgesetzt war.



Bodennahes Ozon und Wetterlage
Obere Kurve: bei lebhaften Westwinden
Untere Kurve: bei schwachen Nordostwinden

Abb. 3.

In Abb. 3 zeigen die oberen Kurven bei lebhaften Westwinden höhere Ozonwerte und Fortfall der nächtlichen Minima. Die Ozonzerstörung am Boden wird dann vollkommen überdeckt durch die Herbeiführung von Ozon aus höheren Schichten. Die starken Schwankungen sind offenbar mit einer etwas großräumigen Turbulenz der Luft in Verbindung zu bringen, die Luftmassen verschiedenen Ozongehaltes, am Bodensee auch Land- und Seeluft an die Meßstelle bringt.

VOL. 2 1952

Bei ruhigem Wetter ist am Bodensee-Ufer auch ein antiparalleler Gang zwischen der Turbulenz und der Feuchtigkeit zu beobachten [4]. Man kann daher die Abnahme der Turbulenz an der Zunahme der Feuchtigkeit abschätzen. Wenn die Luft ruht, dann nimmt sie bald sehr hohe relative Feuchtigkeit an. Wenn dagegen Thermik einsetzt, kommt trockenere Luft aus höheren Schichten herab. Man sieht daher auf Abb. 4 einen antiparallelen Gang zwischen der Feuchtigkeit (die reziprok mit der Turbulenz ist) und dem Ozongehalt. Interessant ist an dieser Kurve noch die plötzliche Erhöhung des Ozongehaltes um $^1/_2$9 Uhr abends. Hier zog eine Böe mit maximal 8 m/sec über den Beobachtungsort. Gleichzeitig mit dem Ozon wurde dabei auch trockenere Luft aus der Höhe herabgeführt, wie an der gleichzeitig auftretenden Zacke an der Feuchtigkeitskurve zu erkennen ist. Als die Böe vorüber war, ging der Ozongehalt auch bald auf Null zurück und die Feuchtigkeit herauf.

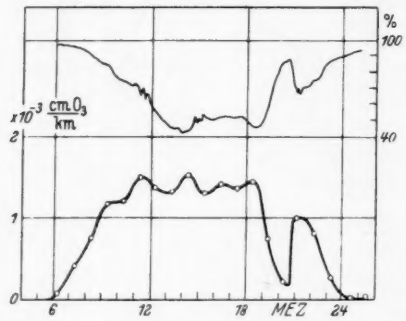


Abb. 4. Ozon (oben) und Luftfeuchtigkeit (unten).

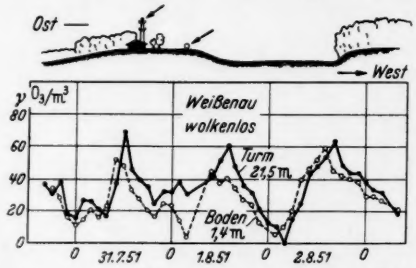


Abb. 5. Ozongang in 1,4 und 20 m Höhe.

Natürlich kann bei einer einsetzenden größeren Turbulenz auch der Ozongehalt am Boden absinken, wenn die allgemeine Wetterlage in der Höhe eine Luftschicht mit geringerem Ozongehalt herangeführt hat als in Bodennähe.

Nicht immer sind die lokalen Verhältnisse so einfach wie am Bodensee-Ufer. Mit der Änderung der Geländekonfiguration müssen sich auch veränderte Tagesgänge des bodennahen Ozons ausbilden, die davon abhängen, wie weit sich Temperaturinversionen und stagnierende Luft bilden können, in der das Ozon am Boden zerstört wird und inwiefern der Luftaustausch durch die Geländegestaltung und durch die lokalen Windverhältnisse beeinflußt wird. Das zeigen insbesondere die vielen Ozonmessungen, die in Badeorten aus klimatologischem Interesse heraus gemacht worden sind. Es kommt besonders an Orten im Gebirge auch vor, daß am Abend oder in der Nacht ein zweites Ozonmaximum auftritt, das durch Fallwinde erklärt werden kann, die ozonreichere Luft herabführen. Überall dort aber, wo es in Niederungen und Tälern in der Nacht zur Ausbildung von Kälteseen mit Temperaturinversionen kommt, tritt das nächtliche Ozonminimum auf.

Um festzustellen, wie hoch die ozonzerstörende Wirkung des Bodens in einem Einzelfalle heraufreicht, haben wir in unserer neuen Forschungsstelle in Weißenau (20 km vom Bodensee entfernt) zwei Meßvorrichtungen nach dem neuen Verfahren von A. Ehmert aufgestellt, eine auf einer Wiese, eine auf einem 20 m hohen Gittermast (Abb. 5). Obgleich das Gelände hier nur mäßig bewegt ist, hat sich gezeigt, daß die täglichen Ozongänge auch bei ruhigem Wetter nicht so einfach sind, wie am Bodensee-Ufer in Friedrichshafen. Das leicht eingeschnittene Gelände, der Wald und ein nahes Flußtal machen ihre Einflüsse auf den Wind und die Turbulenz und

damit auf den täglichen Ozongang geltend. Immerhin lassen sich Tage herausfinden, bei denen der Einfluß der Höhendifferenz von 20 m zwischen den beiden Apparaten deutlich herauskommt. So sieht man in Abb. 5 die Tagesspitzen des Ozons bei höheren Werten liegen und in der Nacht geht auf dem Turm die Ozonkonzentration meist nicht so weit herunter wie am Fuße des Turms.

Mit dem neuen Ozonmeßgerät von EHMERT laufen jetzt an vier Orten in Süddeutschland und einer Stelle in der Schweiz (Arosa) dauernde Ozonregistrierungen, über die A. EHMERT in dieser Zeitschrift berichten wird.

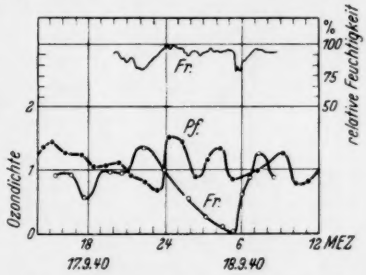


Abb. 6. Ozongang auf dem Pfändergipfel und in Friedrichshafen.

Der Einfluß größerer Höhendifferenzen ist auf Abb. 6 zu erkennen, auf denen der von A. EHMERT gemessene tägliche Ozongang in Friedrichshafen in 400 m Höhe und gleichzeitig auf dem am östlichen Ende des Bodensees in 1064 m gelegenen Pfändergipfel dargestellt wird. Man sieht, daß das nächtliche Absinken des Ozongehaltes, das in Friedrichshafen beobachtet wird, in der bewegteren Luft auf dem nach Westen zu exponiert liegenden Pfändergipfel fehlt. Gelegentlich, so am 17. 9. 1940 um 22 Uhr wird in Friedrichshafen auch mehr Ozon gefunden, was wieder auf eine Umwälzung der Luftmassen hindeutet.

VOL.
2
1952

Noch deutlicher zeigen sich die Unterschiede im täglichen Ozongang bei größeren räumlichen Verhältnissen zwischen einer Tal- und einer Bergstation bei zwei Messungen nach einer chemischen Methode, die kürzlich in New Mexiko (USA.) ausgeführt worden sind (J. G. BOWEN und V. H. REGENER [5]. Die Messungen wurden gleichzeitig auf einem ziemlich isoliert stehenden Gipfel, dem Capillo Peak (2800 m hoch, 900 bis 1500 m über der Umgebung) und einer in einem breiten Tale gelegenen, immerhin noch 2000 m hohen Station Acomita ausgeführt. Abb. 7 zeigt, daß auf der Bergstation nur unregelmäßige, durch den Wetterablauf verursachte Schwankungen des Ozongehaltes auftreten, während in der Talstation regelmäßig starke nächtliche Ozonminima entstehen (die Messungen sind im Dezember gemacht). In der Tat bildeten sich, wie die Temperaturkurve in Abb. 8 zeigt, auf der Acomita-Station regelmäßig auch starke nächtliche Temperaturminima aus, die den Luftaustausch mit der Höhe unterbinden, während dies auf der Bergstation nicht der Fall ist (Abb. 9). Diese Messungen zeigen, daß auch in gebirgigem Gelände in einem weiten Tale das nächtliche Absinken des Ozongehaltes auftritt, allerdings bei den größeren räumlichen Ausmaßen des Geländes auch entsprechend langsam, so daß das Ozonminimum erst gegen Morgen eintritt, am dritten Beobachtungstage sogar auf den nächsten Nachmittag verschoben erscheint. Die desozonisierende Wirkung

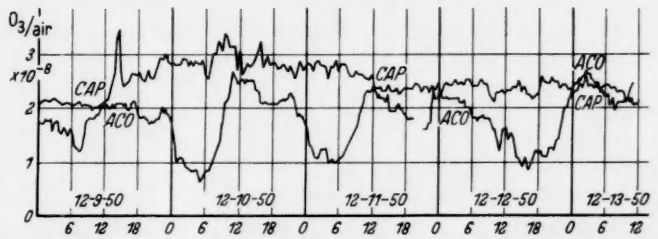


Abb. 7. Ozongang auf einem Berggipfel (*CAP*) und in einem weiten Tal (*ACO*).

des Erdbodens hat sich offenbar unter den gegebenen örtlichen Verhältnissen auf ein großes Luftvolumen erstreckt; der Ausgleich durch die Turbulenz hat offenbar

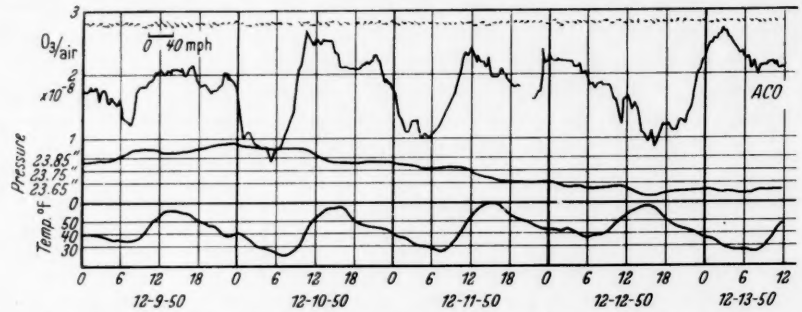


Abb. 8. Ozongang, Luftdruck und Temperatur in einem weiten Tale (*ACO*).

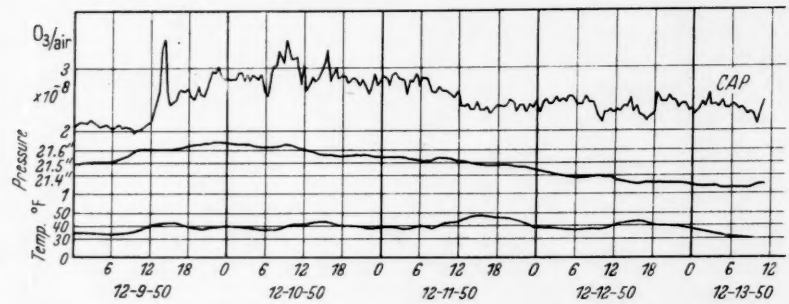
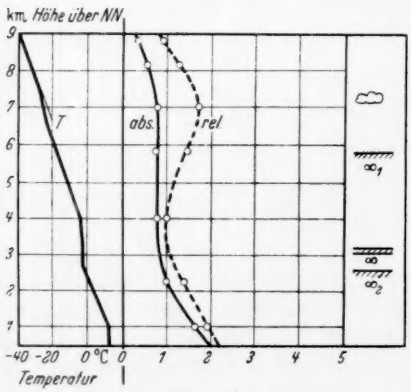
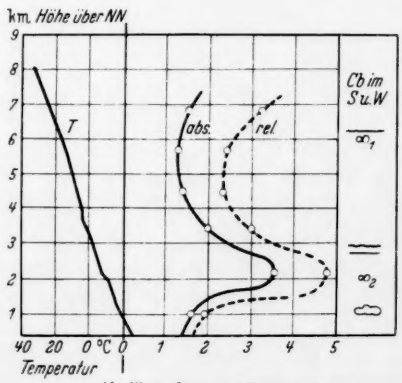


Abb. 9. Ozongang, Luftdruck und Temperatur auf einem Berggipfel (*CAP*).



Vertikale Ozonverteilung am 17.8.42
Nach A. Ehmert
abs. = 10^{-3} cm. Normalozon/km. Luftweg
rel. = 10^{-8} Ozon/Luft

Abb. 10.

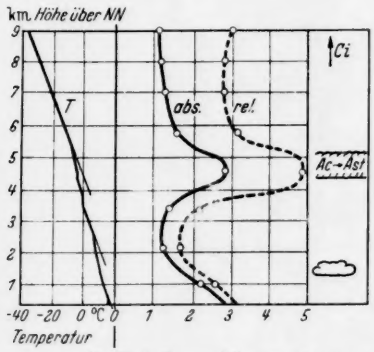


Vertikale Ozonverteilung am 19.8.42
Nach A. Ehmert
abs. = 10^{-3} cm. Normalozon/km. Luftweg
rel. = 10^{-8} Ozon/Luft

Abb. 11.

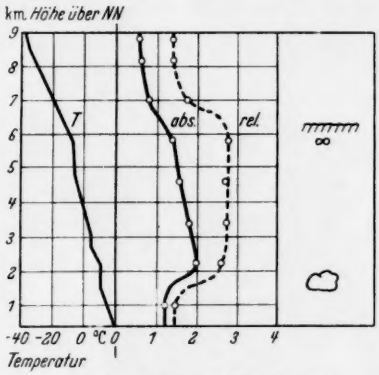
länger gedauert. Allgemein wird man sagen können, daß die ozonfreien und ozonarmen Gebiete um so größer und um so persistenter sein müssen, je günstiger die örtlichen Verhältnisse für die Ausbildung großräumiger Kälteseen in der Nacht liegen.

Was nun die schwankende Verteilung des Ozons in der freien Troposphäre betrifft, so kann ich dafür nur auf Messungen von A. EHMERT [6] zurückgreifen, die er noch im Kriege nach der chemischen Methode im Flugzeug gemacht hat. Die vier Messungen in Abb. 10–13 liegen je 2–3 Tage auseinander und reichen mit einer Ausnahme bis 8·5 und 9 km Höhe. Am ersten Tage (Abb. 10) sieht man nach einem Minimum in 4 km Höhe einen schwachen Anstieg des Ozon/Luft-Verhältnisses in 7 km Höhe. Nach zwei weiteren Tagen (Abb. 11) zeigt sich in 2 km Höhe ein ziemlich starkes Maximum, das nach zwei weiteren Tagen auf 4·5 km Höhe gerückt ist (Abb. 12). Drei Tage später (Abb. 13) ist das Maximum stark verwischt, dafür



Vertikale Ozonverteilung am 21.8.42
Nach A. Ehmert
abs.=10^{-3} cm Normalozon/km Luftweg
rel.=10^{-8} Ozon/Luft

Abb. 12.



Vertikale Ozonverteilung am 24.8.42
Nach A. Ehmert
abs.=10^{-3} cm Normalozon/km Luftweg
rel.=10^{-8} Ozon/Luft

Abb. 13.

zeigt aber die Kurve für das Verhältnis Ozon/Luft zwischen 2 und fast 6 km denselben Wert, zum Zeichen dafür, daß in der dazwischen liegenden Luftschicht eine starke Durchmischung stattgefunden hat. Die daneben gezeichnete Temperaturkurve zeigt ebenso wie die angedeuteten Wolken die Abgrenzung dieser Schicht durch Inversionen. Es kann nach den gezeigten Kurven wohl kein Zweifel darüber bestehen, daß für die Ozonverteilung in der Troposphäre sowohl die Advektion wie die Turbulenz, der vertikale Luftaustausch eine Rolle spielen. Welcher Faktor im Einzelfalle überwiegt, das kann wohl erst entschieden werden, wenn die vertikale Ozonverteilung in der Troposphäre an mehreren Orten gleichzeitig bestimmt sein wird.

Was schließlich die Ozonverteilung in noch größeren Höhen betrifft, so zeige ich dafür zunächst eine Zusammenstellung von acht Verteilungen (Abb. 14), alten und neuen, die bis auf eine Ausnahme [8] bei Ballonaufstiegen mit Spektrographen gewonnen sind. Es überrascht die Verschiedenartigkeit der Kurven gegenüber unserer Kenntnis vor 10 oder mehr Jahren. Neben den bekannten älteren Kurven mit einem Maximum in der Gegend bei 20 oder mehr km Höhe sieht man jetzt Kurven mit tieferer und höherer Lage des Maximums und auch solche, die ein zweites Maximum in tieferer Lage zeigen, wie solches auch schon nach der Umkehrmethode [7] gefunden worden ist. Ganz besonders bemerkenswert ist, daß das Verhältnis Ozon/Luft in Höhen über 20 km besonders stark variiert. Dies kann nur so erklärt werden, daß in dieser Höhe eine starke und schnelle Durchmischung eingetreten sein

muß. Die zweiten Maxima in geringeren Höhen, die meist im Frühjahr gefunden werden, sind wohl in der Hauptsache durch Advektion von niedrig liegenden polaren Luftkörpern mit hohem Ozongehalt zu deuten. Mit der photochemischen Theorie ist wohl nur das Auftreten eines Maximums in etwa 23 km Höhe zu vereinigen wie es bei den ersten früheren Ballonaufstiegen gefunden wurde. Mein Mitarbeiter

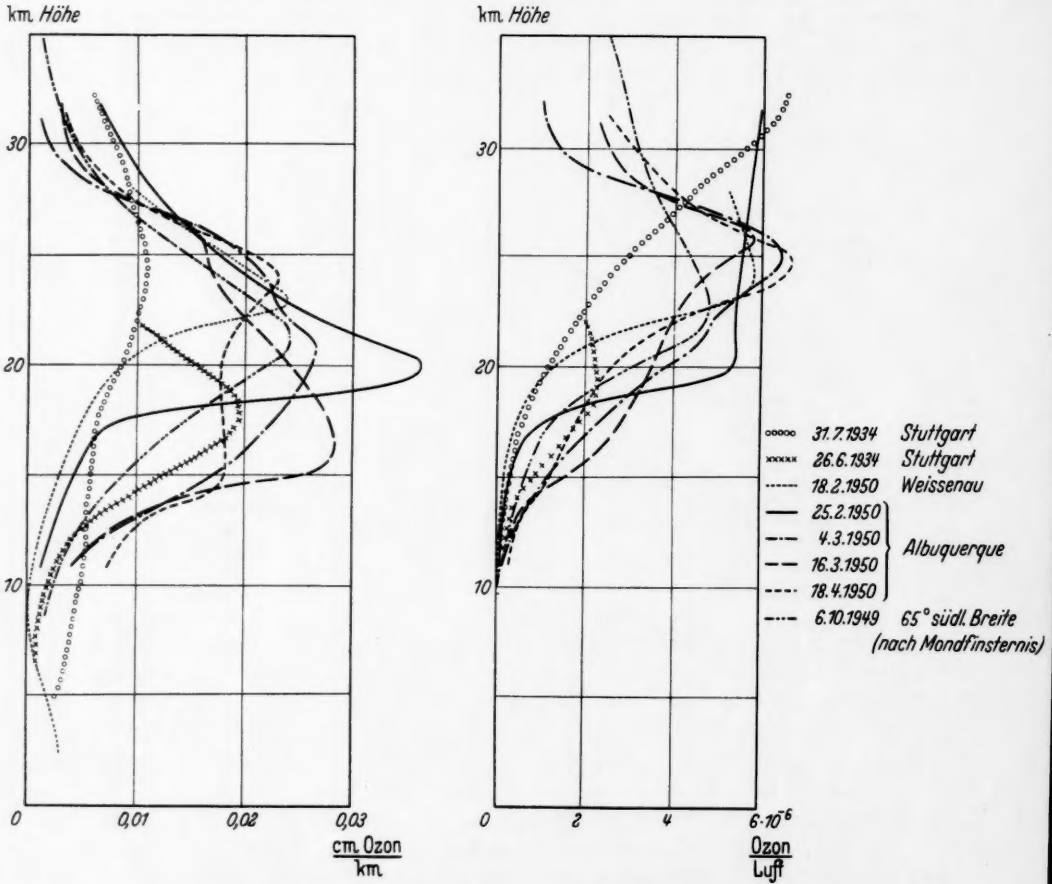


Abb. 14. Stark unterschiedliche vertikale Ozonverteilungen.

H. K. PAETZOLD hat mit den von SCHRÖER [9] gegebenen Ansätzen die vertikale Ozonverteilung noch einmal durchgerechnet, dabei aber die eingehenden Konstanten, wie die (noch unsicher bekannten) Absorptionskoeffizienten des Sauerstoffs, die für die Ozonbildung in Betracht kommen, variiert, ebenso das BEERsche Gesetz entsprechend probeweise abgeändert und auch die Annahme für die Temperatur in den für die Ozonbildung in Betracht kommenden Höhen variiert. Das Resultat (Abb. 15) zeigt aber, daß das alles nicht viel ausmacht, sondern daß die Kurven sehr ähnlich sind und in ihrem Verlauf, wenn auch nicht in der Höhe des Maximums mit denen bei ruhigen Wetterlagen gefundenen (stark ausgezogenen Kurven) übereinstimmen.

Wenn also die beobachteten Kurven anders verlaufen, so muß die Differenz zwischen der beobachteten Kurve als ein Maß für die Einflüsse der Advektion oder Turbulenz gedeutet werden. Auf alle Fälle ergeben die gezeigten Kurven, daß in großen Höhen beachtenswert starke und sonst noch wenig untersuchte horizontale und vertikale Bewegungen vor sich gehen.

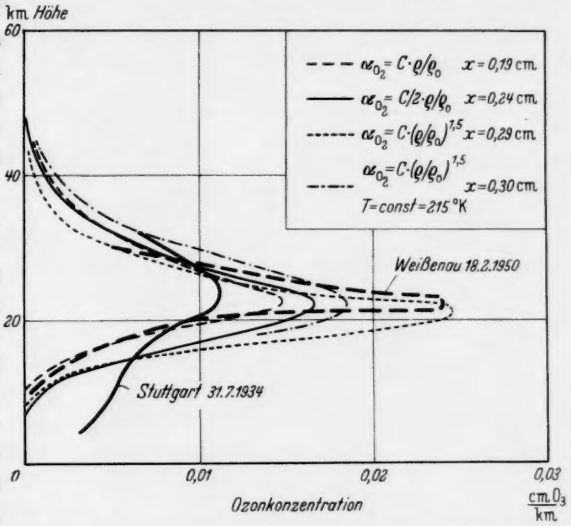
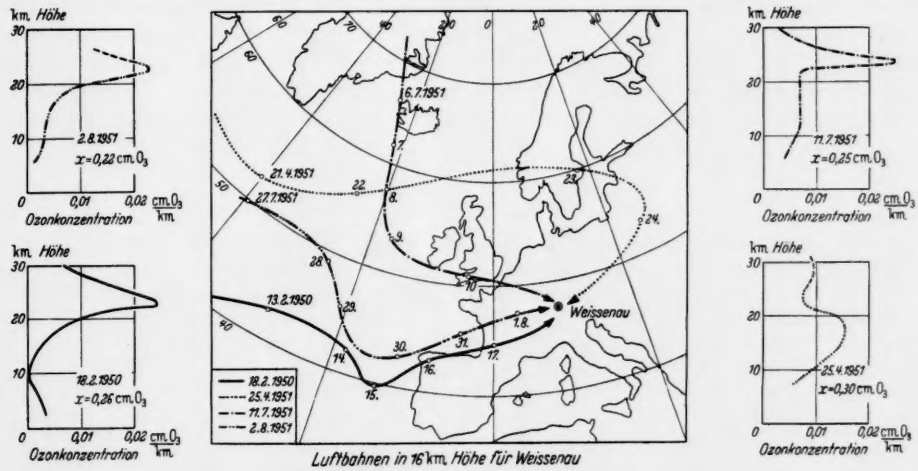


Abb. 15. Theoretische vertikale Ozonverteilungen.

In Abb. 16 ist der von H. K. PAETZOLD gemachte Versuch dargestellt, an vier kürzlich mit dem Spektrographen gewonnenen Ozonkurven, die Form der Kurven in Zusammenhang mit der Herkunft der Luftmassen zu bringen. Die Herkunft der Luftmassen ist natürlich nur mit Vorbehalt anzugeben, denn es sind die Trajektorien der Luft in 16 km Höhe benutzt worden, und wie die Luft sich in größeren Höhen bewegt, ist natürlich unbekannt. Immerhin deutet sich schon ein gewisser Zusammenhang an, indem bei den beiden linken Ozonverteilungskurven, die ein spitzes



Vertikale Ozonverteilungen und Luftbahnen

Abb. 16.

Maximum zeigen, Luft aus der Azorengegend herangeführt wird, während für die Verteilungskurve rechts unten, die ein starkes Maximum unterhalb von 20 km und die Andeutung eines zweiten in größerer Höhe zeigt, die Trajektorie nach

VOL. 2 1952

Nordamerika weist. Die obere rechte Kurve, bei der die Luft in 16 km Höhe von Grönland kommt, zeigt neben einem steilen Maximum in 23 km Höhe immerhin zwischen 10 und 20 km Höhe ziemlich konstanten und einigermaßen hohen Ozongehalt. Die dem photochemischen Gleichgewicht am ehesten entsprechende Kurve mit einem spitzen Maximum, die auch meist mit einem geringeren Gesamtwert des Ozons verbunden ist, scheint also den äquatorialen Luftmassen zu entsprechen, die deformierten Kurven den bei uns meist im Frühjahr auftretenden polaren Luftmassen mit einem höheren Gesamtgehalt an Ozon [10].

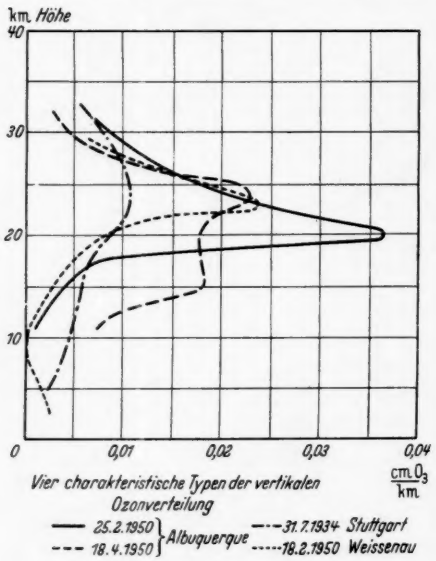


Abb. 17.

Zur besseren Übersicht sind in Abb. 17 noch einmal vier charakteristische mit dem Spektrographen erhaltene Fälle der vertikalen Ozonverteilung zusammengestellt [11], [12].

ZUSAMMENFASSUNG

Die ganze Kompliziertheit des Problems des atmosphärischen Ozons spiegelt sich in den Schwankungen der Ozonkonzentration in allen Höhen wieder. In der unmittelbar auf dem Boden liegenden Luftschicht sinkt infolge der starken ozonzerstörenden Wirkung des Erdbodens der Ozongehalt immer dann auf Null, wenn die Luft stagniert. In der Troposphäre spielt die Advektion die Hauptrolle bei den Ozonschwankungen. In größeren Höhen sind mit der spektrographischen Methode bei Ballonaufstiegen neuerdings stark unterschiedliche vertikale Ozonverteilungen gefunden worden. Für die Erklärung müssen großräumige horizontale und vertikale Luftbewegungen in großen Höhen angenommen werden.

REFERENCES

[1] MOSER; Ber. Dtsch. Wetterdienst. US-Zone (Nr. 11) p. 28, 1949. [2] AUER, R.; Gerl. Beitr. Geophys. 1939 **54** 137. [3] EHMERT, A.; Ber. Dtsch. Wetterdienst. US-Zone (Nr. 11) p. 47, 1949. [4] EHMERT, A.; *loc. cit.* [3]. [5] BOWEN, J. G. und REGENER, V. H.; J. Geophys. Res., September 1951. [6] EHMERT, A.; *loc. cit.* [3], p. 63. [7] GÖTZ, F. W. P.; Gerl. Beitr. Geophys. 1931 **31** 119. [8] Kurve vom 6. 10. 1949 bei PAETZOLD, H. K.; Z. f. Naturforschung 1950 **5a** 661. [9] SCHRÖER, E.; Ber. Dtsch. Wetterdienst. US-Zone (Nr. 11) p. 13, 1940. [10] Vgl. hierzu MOSER; *loc. cit.* [1]. [11] REGENER, E., PAETZOLD, H. K. und PFOTZER, G.; Naturwiss. 1950 **37** 559. [12] REGENER, V. H.; Nature, Lond. 1951 **167** 276.

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 183 to 188. Pergamon Press Ltd., London

Erfassung der vertikalen Ozonverteilung in verschiedenen geographischen Breiten bei Mondfinsternissen*

Determination of Vertical Ozone Distribution at Various Latitudes During Eclipses of the Moon

H. K. Paetzold

Forschungsstelle Weißenau

(Received 20 September 1951)

ABSTRACT

The vertical distribution of atmospheric ozone may be determined from the distribution of intensity, measured at 6000 Å, on the darkened moon, near the edge of the earth's shadow. It is a characteristic of the method that the ozone distribution at various latitudes may be determined regardless of the point of observation. The usefulness of the method is demonstrated by spectrophotometric measurements made during four eclipses of the moon. The ozone distribution thus obtained correspond to results obtained directly by balloon ascents.

VOL. 2 1952

Bekanntlich wird bei einer Mondfinsternis Sonnenlicht, das durch die Erdatmosphäre läuft, in den geometrischen Kernschattenraum hineingebrochen. Dabei durchläuft das Licht um so tiefere Luftchichten, je weiter es in den Schattenraum auf der Mondoberfläche eindringt, d. h. je größer die Entfernung γ' eines von ihm beleuchteten Punktes von der geometrischen Schattengrenze ist. Eine in die Erdatmosphäre eingelagerte, spektral absorbierende Schicht und ihre Höhenverteilung müssen sich daher aus ihrer in Abhängigkeit von γ' gemessenen Absorption bestimmen lassen. Auf diese Weise kann man die vertikale Ozonverteilung für verschiedene geographische Breiten der Erde bestimmen, je nach dem Positionswinkel, unter dem die Beleuchtungsverteilung auf dem verfinsterten Mond senkrecht zur Schattengrenze gemessen wird. Der erfaßbare Positionswinkel hängt dabei von der Bahn des Mondes durch den Erdschatten ab, ist aber von der geographischen Lage des Beobachtungsortes auf der Erde unabhängig.

Um auf diese Weise die Ozonabsorption bestimmen zu können, muß die bei einer ozonfreien Atmosphäre herrschende Beleuchtungsverteilung am Rande des Erdschattens auf dem Mond bekannt sein. Diese kann nach einer von F. Link [1] entwickelten photometrischen Theorie der Mondfinsternisse berechnet werden. Nach eigenen Messungen [2] ist in den von der Ozonabsorption freien Spektralgebieten die Übereinstimmung zwischen Beobachtung und Theorie so ausreichend, daß letztere als Grundlagen für die Ozonmessung dienen kann. Bei diesen Beobachtungen und denen anderer Autoren [1], [3], [4] zeigt sich in der Tat im Gebiet der Chappuis-Banden des Ozons eindeutig eine zusätzliche Schwächung des Lichtes bis auf $^1/_{10}$, die nur dem Ozon zugeschrieben werden kann††.

* Vorgelegt auf dem Ozonsymposion in Brüssel am 30. 8.–1. 9. 1951 durch E. Regener.

** Diese starke Absorption ist, wie Verf. zeigen konnte [5], auch der Grund für eine charakteristische grünliche Verfärbung einer schmalen Zone längs des Schattenrandes. Durch diese „grüne Zone" am Rande des Erdschattens auf dem Mond ist die atmosphärische Ozonschicht direkt mit dem Auge wahrzunehmen.

Theoretisch muß sich also hieraus die vertikale Ozonverteilung für die auf der Mondoberfläche erfaßbaren Breiten der Erde bestimmen lassen. Die bisherigen in dieser Richtung unternommenen Versuche [1], [3] sind aber über Ansätze nicht hinausgekommen. Denn die Auswertung der Beobachtungen wird durch die Tatsache erschwert, daß die Sonne kein punktförmiger Strahler ist, sondern von Erde und Mond aus als Scheibe von 32′ Durchmesser erscheint. Daher wird ein im Schattenraum liegender Punkt der Mondoberfläche nicht von einem Lichtstrahl, sondern von einem Lichtbündel von beträchtlichem Querschnitt beleuchtet, wodurch natürlich in bezug auf die zu berechnende Ozonverteilung eine gewisse Verschmierung eintreten muß. Zur Abschätzung dieses letzteren Einflusses wurde für drei vorgegebene Ozonverteilungen $\varepsilon(h)$ die mittlere Ozonmasse $O^*(\gamma')$† berechnet, die ein Lichtbündel durchläuft.

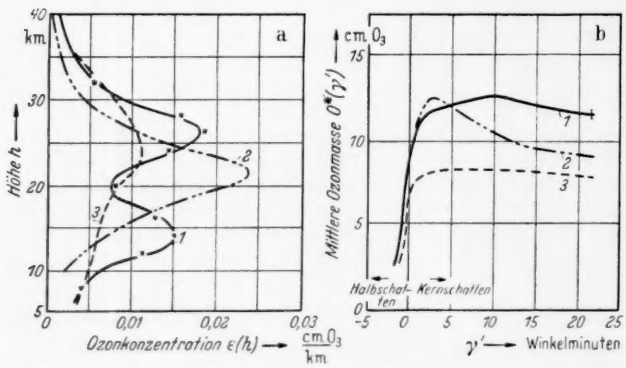


Abb. 1a u. b. a Ozonverteilungen. *1* GOETZ [6], Umkehreffekt; *2* PAETZOLD [2], Mondfinsternis, 6. Oktober 1949; *3* E. und V. H. REGENER [7], 31. Juli 1934. b Mittlere Ozonmassen $O^*(\gamma')$, die von dem Lichtbündel bei den verschiedenen Ozonverteilungen in Abb. 1a durchlaufen werden (gleiche Numerierung)††.

In Abb. 1 sind links diese drei Ozonverteilungen vorgezeichnet, die sich in ihrem Charakter stark voneinander unterscheiden, nämlich flaches, steiles und doppeltes Maximum. Die dazugehörigen Kurven in Abb. 1 rechts zeigen die Abhängigkeit der mittleren Ozonmassen $O^*(\gamma')$ von der Entfernung γ' des jeweils von dem Lichtbündel beleuchteten Punktes des Mondes von dem geometrischen Schattenrand ($\gamma'=0$). Wie man sieht, ist trotz der erwähnten Verschmierung durch die Ausdehnung des Lichtbündels der Verlauf der Kurven so verschieden, daß sich daraus der allgemeine Verlauf der vertikalen Ozonverteilung sollte entnehmen lassen.

Nun ist freilich die tatsächlich vorzunehmende Ableitung der unbekannten Ozonverteilung $\varepsilon(h)$ aus der gemessenen Funktion $O^*(\gamma')$ in geschlossener analytischer Form nicht möglich. Sie würde gelöst sein, wenn die Ozonmasse $O(h_0)$ bekannt ist, die ein Einzelstrahl durchläuft, der die Atmosphäre in dem Minimalabstand h_0 von der Erdoberfläche tangential passiert (Abb. 2). Denn bei bekannter Ozonverteilung $\varepsilon(h)$ ist diese Ozonmasse $O(h_0)$ gegeben durch:

$$O(h_0) = \int_{-\infty}^{+\infty} \varepsilon(h)\, ds, \tag{1}$$

worin ds ein Wegelement des in Abb. 2 von B nach A verlaufenden Einzelstrahles bedeutet.

† Alle sich auf das Lichtbündel beziehenden Größen werden im folgenden mit * bezeichnet.

†† Der größte Unterschied zwischen der Kurve *1* und *3* entspricht beim Maximum der CHAPPUIS-Bande (6000 Å) intensitätsmäßig dem Faktor 2.

Durch Einsetzen des aus dem gezeichneten Dreieck in Abb. 2 folgenden Wertes für ds erhält man:

$$O(h_0) = \sqrt{2a} \int_{h_0}^{\infty} \frac{\varepsilon(h)\, dh}{\sqrt{h - h_0}} \cdot \dagger \tag{2}$$

Ist, wie im vorliegenden Fall, $\varepsilon(h)$ die gesuchte Funktion, so stellt Gl. 2 eine ABELsche Integralgleichung für $\varepsilon(h)$ dar, deren Lösung lautet:

$$\varepsilon(h) = -\frac{1}{\pi\sqrt{2a}} \int_{h}^{\infty} \frac{\left(\frac{dO(h_0)}{dh_0}\right)}{\sqrt{h_0 - h}}\, dh_0. \tag{3}$$

Die Ableitung von $O(h_0)$ aus der aus den Beobachtungen folgenden mittleren Ozonmasse $O^*(\gamma')$ gelingt nun verhältnismäßig einfach durch folgenden Kunstgriff: Man führt in $O^*(\gamma')$ statt der Variablen γ' den Minimalabstand h^* von der Erdoberfläche desjenigen Einzelstrahles (Ersatzstrahles) ein, der dieselbe Luftmasse durchläuft wie das Lichtbündel im Mittel, so daß man für $O^*(\gamma')$ die neue Funktion $O^*(h^*)$ erhält. Diese Substitution von γ' wird durch die Tatsache nahegelegt, daß für kleinere Höhen die Funktionen $O(h_0)$ und $O^*(h^*)$ sich mehr und mehr annähern. Denn für kleinere Höhen passiert das gesamte Lichtbündel die Ozonschicht unterhalb ihres Maximums. Die Rechnung zeigt für diesen Fall, daß die von den Einzelstrahlen des Lichtbündels durchlaufenen Ozonmassen sich im Gegensatz zu größeren Höhen verhältnismäßig wenig voneinander unterscheiden. Aus diesem Grund ist die von dem Einzelstrahl in der Mitte des Lichtbündels durchlaufene Ozonmasse sehr angenähert gleich der mittleren Ozonmasse $O^*(\gamma')$. Dasselbe gilt auch in bezug auf die RAYLEIGH-Extinktion. Dieser Effekt wird ferner durch die Tatsache unterstützt, daß mit kleiner werdenden Höhen die Höhenausdehnung des Lichtbündels stark abnimmt, so daß letzteres sich dem Einzelstrahl annähert. Abb. 3 erläutert diese Verhältnisse für die drei Ozonverteilungen nach Abb. 1 links.

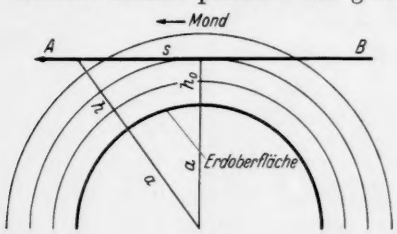


Abb. 2.

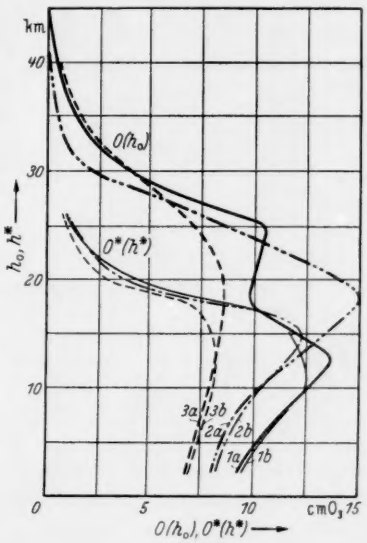


Abb. 3.

Abb. 2. Bahn eines Einzelstrahles in der Atmosphäre (schematisch).

Abb. 3. Die Kurven *1a, 2a, 3a* zeigen den Verlauf der von einem Einzelstrahl durchlaufenen Ozonmasse $O(h_0)$ für die Ozonmasse *1, 2, 3* in Abb. 1 links, während die Kurven *1b, 2b, 3b* die von dem Ersatzstrahl durchlaufene Ozonmasse $O^*(h^*)$ geben. Man beachte das fast völlige Zusammenfallen der Kurven $O(h_0)$ und $O^*(h^*)$ unterhalb von etwa 10 km Höhe.

† Die Refraktion kann hier vernachlässigt werden.

VOL. 2 1952

Unterhalb von 10 km fallen die vom Ersatzstrahl durchlaufenen Ozonmassen $O^*(h^*)$ mit der gesuchten Funktion $O(h_0)$ praktisch zusammen, während sich beide Funktionen für größere Höhen immer stärker voneinander unterscheiden. Daher gelingt es, von kleineren zu größeren Höhen fortschreitend $O(h_0)$ schrittweise aus $O^*(h^*)$ abzuleiten, indem die Extrapolation durch Rückwärtsrechnung jeweils geprüft und verbessert wird.

Um die Sicherheit des Verfahrens zu erproben, wurde aus den berechneten mittleren Ozonmassen in der Kurve 1, Abb. 1 links, die dazugehörige Ozonverteilung nach der obigen Methode rückwärts abgeleitet. Die erhaltenen, mit Sternen in Abb. 1 links eingetragenen Kontrollpunkte liegen, wie man sieht, recht gut auf der angenommenen Verteilungskurve. Selbstverständlich müssen zu feine Einzelheiten einer Verteilung durch die oben erwähnte Verschmierung verlorengehen.

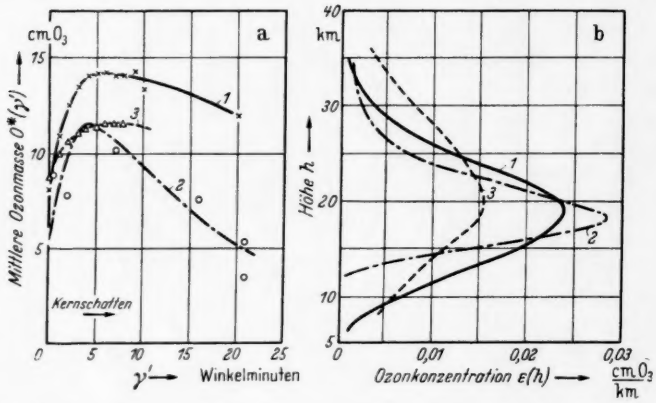


Abb. 4a u. b. a Mittlere Ozonmassen $O^*(\gamma')$ nach spektralphotometrischen Messungen auf der Oberfläche des verfinsterten Mondes. *1* 16. Oktober 1921 [3]; *2* 2. März 1942 [4]; *3* 26. September 1931 [1]. b Vertikale Ozonverteilungen nach den mittleren Ozonmassen in a.

Eine nach der vorliegenden Methode ausgewertete Messung bei der Mondfinsternis am 6./7. Oktober 1942 [2] ergab oberhalb von 45 km Höhe eine verschwindende Ozonmenge, was den herrschenden theoretischen Vorstellungen über die atmosphärische Ozonschicht entspricht und auch bei den amerikanischen Messungen mit der V 2-Rakete gefunden wurde.

Spektralphotometrische Messungen, die nach der geschilderten Methode für die Ozonverteilung ausgewertet werden können, liegen außerdem von LINK [1], DANJON [3] und BARBIER u. a. [4] vor. Die daraus folgenden mittleren Ozonmassen $O^*(\gamma')$ zeigt Abb. 4 links. Die dazugehörigen nach dem geschilderten Verfahren abgeleiteten Ozonverteilungen sind in Abb. 4 rechts enthalten.

Tabelle 1 gibt noch die Daten für den gesamten Ozonbetrag (Ozonwert), für die Höhe des Ozonmaximums und, was für die geschilderte Methode besonders charakteristisch ist, für die geographische Breite, auf die sich gemäß dem Positionswinkel auf dem Mond die Messungen beziehen. Nr. 4 bezieht sich auf den vom Verfasser bereits ausgewerteten Fall [2].

Wie man sieht, entsprechen die Ozonverteilungen, wie sie aus den Mondfinsternissen erhalten wurden, den Ergebnissen der direkten Messungen bei Ballonaufstiegen. Auch der Ozonwert liegt mit 0,22 cm bis 0,33 cm in der Größe der mit anderen Methoden gewonnenen. Er ist nach Tabelle 1 im Mittel für die südpolaren Breiten größer als für die äquatorialen. Zu den Beobachtungszeiten herrschte auf der Südhalbkugel Frühjahr, so daß sich hier ebenfalls die höheren Frühjahrswerte des Ozons

Tabelle 1.

Nr. in Abb. 4	*Datum der Finsternis*	*Geographische Breite*	*Maximale Sonnenhöhe*	*Ozonwert* cm	*Höhe des Ozon-maximums* km
1	16. Oktober 1921	60° S	37°	0·33	19
3	26. September 1931	0°	90°	0·26	20
2	2. März 1942	20° N	63°	0·22	18
4	6. Oktober 1949	65° S	44°	0·25	21

für die höheren Breiten andeuten. Die Höhe des Ozonmaximums zeigt nach Tabelle 1 keine breitenbedingten Unterschiede.

Die vertikale Ozonverteilung, wie sie sich aus dem photochemischen Gleichgewicht berechnen läßt, wird durch vertikalen und horizontalen Austausch stark

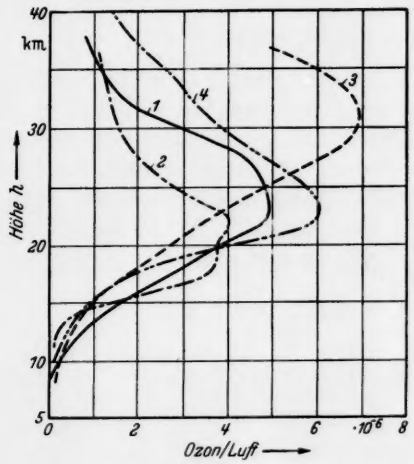


Abb. 5.

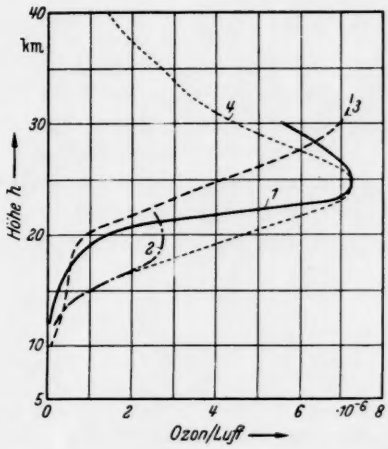


Abb. 6.

Abb. 5. Das Mengenverhältnis Ozon/Luft nach Mondfinsternissen. *1* 16 Oktober 1921; *2* 2. März 1942; *3* 26. September 1931; *4* 6. Oktober 1949.

Abb. 6. Das Mengenverhältnis Ozon/Luft nach Ballonaufstiegen. *1* 18. Februar 1950 [8]; *2* 26. Juni 1934 [8]; *3* 31. Juli 1934 [7]. Kurve *4* wurde zum Vergleich nach dem photochemischen Gleichgewicht für eine Sonnenhöhe von 45° berechnet.

gestört. Der vertikale Austausch sucht dabei ein mit der Höhe konstantes Mengenverhältnis Ozon/Luft herzustellen. Es ist daher in Abb. 5 dieses Mengenverhältnis Ozon/Luft für die vier bei Mondfinsternissen erhaltenen Verteilungen (Tabelle 1) dargestellt und in Abb. 6 dasselbe zum Vergleich für drei Ballonaufstiege in Stuttgart und Weißenau (ca. 48° N) dargestellt. Kurve 4 in Abb. 6 gibt die nach dem photochemischen Gleichgewicht für 45° Sonnenhöhe berechnete Kurve wieder. Wie man sieht, erweisen sich die bei einer geographischen Breite auftretenden Schwankungen des Mengenverhältnisses Ozon/Luft als so groß, daß sie die durch die Breite (Sonnenhöhe) selbst bedingten Variationen vollkommen überdecken.

Die wenigen vorliegenden Messungen zeigen, daß die skizzierte Methode einen guten Beitrag zur Gewinnung eines genügend großen, über alle Breiten der Erde

VOL. 2 1952

verteilten Beobachtungsmaterials der vertikalen Ozonverteilung geben kann. Deshalb seien für die Durchführung der Messungen noch einige ergänzende Hinweise gegeben.

Für die Beobachtungen genügen Farbfilteraufnahmen des verfinsterten Mondes bei 6000 Å, dem Maximum der CHAPPUIS-Banden des Ozons und außerdem zur Kontrolle und eventuellen Eliminierung zusätzlicher Einflüsse (Dunst, zeitweilige abnormale Trübungen in der Stratosphäre usw.) bei 5000 und 4000 Å. Bei diesen Wellenlängen sind Störungen durch die Banden des atmosphärischen Sauerstoffes und durch die sich in größeren Höhen andeutende schwach blauabsorbierende Schicht [1], [2] in 100 km Höhe nicht zu befürchten. Die spektral unterschiedliche Albedo der Mondoberfläche muß natürlich durch Vergleichsaufnahmen des unverfinsterten Vollmondes eliminiert werden. Eine ausführlichere Darstellung der geschilderten Methode und eine nähere Diskussion der Fehlermöglichkeiten ist in der Zeitschrift für Naturforschung erschienen [9].

ZUSAMMENFASSUNG

Aus der Helligkeitsverteilung bei 6000 Å auf dem verfinsterten Mond senkrecht zur Schattengrenze läßt sich die Verteilung des atmosphärischen Ozons ableiten. Für die Methode ist typisch, daß auf diese Weise—unabhängig vom Beobachtungsort—die Ozonverteilung für verschiedene Breiten auf der Erde erfaßt werden kann. An vier bislang bei Mondfinsternissen gewonnenen spektralphotometrischen Messungen wird die Brauchbarkeit des Verfahrens gezeigt. Die so erhaltenen Ozonverteilungen entsprechen den auf direktem Wege bei Ballonaufstiegen gewonnenen Ergebnissen.

LITERATUR

[1] LINK, F.; Bull. Astronom. 1932 **8** 77. [2] PAETZOLD, H. K.; Z. Naturforschung 1950 **5a** 661. [3] DANJON, C. R.; 1921 **173** 706. [4] BARBIER, D., CHALONGE, D. und VIGRAUX, E.; Ann. Astrophys. 1942 **5** 1. [5] PAETZOLD, H.K.; Naturwiss. 1951 **38** 544. [6] GOETZ, F.W.P.; Ber. Dtsch. Wetterdienst US-Zone (Nr. 11) p. 7, 1949. [7] REGENER, E. und V.H.; Phys. Z. 1934 **35** 788. [8] REGENER, E., PAETZOLD, H. K. und PFOTZER, G.; Naturwiss. 1950 **37** 559. [9] PAETZOLD, H. K.; Z. Naturforschung 1951 **5a** 639.

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 189 to 195. Pergamon Press Ltd., London

Gleichzeitige Messungen des Ozongehaltes bodennaher Luft an mehreren Stationen mit einem einfachen absoluten Verfahren *

Simultaneous Measurements of the Ozone Content of Air near the Earth at Several Stations by Means of a Simple Absolute Method

Alfred Ehmert

Forschungsstelle für Physik der Stratosphäre (in der Max-Planck-Gesellschaft), Weißenau

(Received 20 September 1951)

ABSTRACT

The local ozone concentration of atmospheric air has been measured over a period at several stations by means of a chemical method which is simple in operation. Several characteristic series of results are given, clearly illustrating the destruction of ozone in the ground layer and the dominating influence of the exchange of air with higher layers. The method of measurement is briefly described.

VOL. 2 1952

Zur Messung der lokalen Ozonkonzentration in der Atmosphäre sind heute die chemischen Methoden der optischen Methode überlegen. Unser unten beschriebenes Verfahren zeichnet sich durch besondere Einfachheit der Handhabung aus, weil die umständliche Verwendung von Normallösungen durch die mühelose Messung von Strömen und deren Einschaltdauer ersetzt wird und weil außerdem schon mit geringen Luftmengen eine ausreichende Genauigkeit erzielt wird. Dies ermöglicht rasches Arbeiten. Die Handhabung der entwickelten Geräte setzt keine chemischen Spezialerfahrungen oder besondere Fertigkeiten voraus. Außerdem gehen in das Resultat nur leicht kontrollierbare Eichgrößen ein, so daß man mit verschiedenen Geräten unmittelbar vergleichbare Absolutwerte erhält.

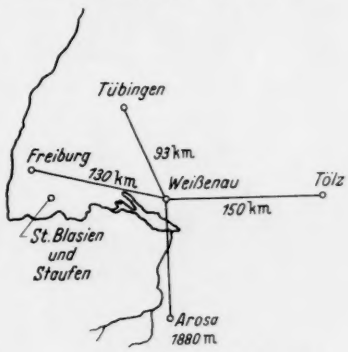


Abb. 1.

Diese Umstände ermutigten uns, mehrere Stationen mit solchen Geräten zu versehen und so eine weiträumige Überwachung des bodennahen Ozons durchzuführen. Abb. 1 zeigt die Lage der Meßorte, von denen Arosa, wo Herr Götz und Herr Volz die Messungen ausführen, durch seine Höhe von 1900 m ausgezeichnet ist. In Tübingen führt die Messungen Herr Daubert im Wetterdienst von Württemberg-Hohenzollern aus. In Freiburg mißt Herr Person (Baldenweger Hof) beim Südbadischen Wetterdienst und im selben Dienst Herr Busse in St. Blasien bzw. Staufen im Schwarzwald. In Bad Tölz mißt Herr Ungeheuer an der dortigen Klimastelle des Wetterdienstes der US-Zone.

Genaue meteorologische Auswertungen werden an anderer Stelle veröffentlicht. Hier werden aus dem Material einige charakteristische Fälle ausgewählt. Den an den Meßreihen beteiligten genannten Herren danke ich für die freundlichst zur Verfügung gestellten Meßergebnisse. Die tatsächlichen Werte von Arosa sind zum

* Vorgetragen auf dem Ozon-Symposium der *JGGU* in Brüssel, am 30. 8. 1951.

Vergleich auf den Luftdruck der anderen Stationen mit einem mittleren Faktor umgerechnet.

In Abb. 2 liegt zunächst zyklonale Druckverteilung mit böigen Westwinden vor. In Arosa sind die Ozonwerte um 20% höher. Am 28. 4. 1950 flauten die Winde ab und in Tübingen sank der Ozongehalt ab. Die dortige Meßstelle befindet sich im Schloß etwa 20 m über den Dächern von Tübingen bei freiem Zutritt westlicher Winde.

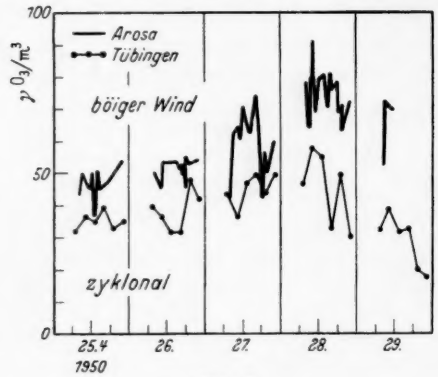


Abb. 2.

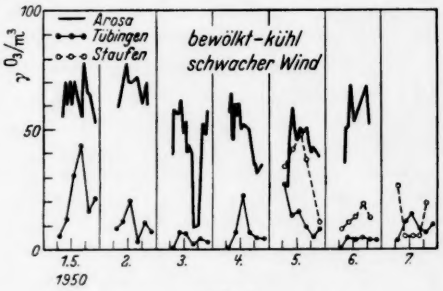


Abb. 3.

Abb. 3 zeigt den anschließenden Zeitraum mit schwachen Winden bei bedecktem Himmel und kühlem Wetter in Tübingen. Der Luftkörper der Stadt Tübingen hat viel Ozon verloren und die Zufuhr frischer Luft aus der Höhe ist nur sehr gering. In Arosa sind die Bodeneinflüsse weit geringer und treten nur stärker hervor, wenn Luft aus dem Tal über die Stadt an die Meßstelle transportiert wird, wie z. B. am 3. 5. 1950.

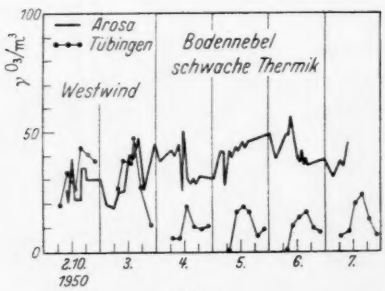


Abb. 4.

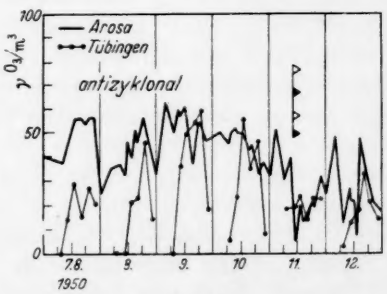


Abb. 5.

Abb. 4 zeigt einen ähnlichen Übergang von Westwind zu stagnierenden Verhältnissen in Tübingen. Die herbstliche schwache Thermik reicht nicht zu einem vollen Anschluß an die vom Boden unbeeinflußten Schichten aus.

Dagegen zeigt Abb. 5, wie dieser Anschluß durch die starke sommerliche Thermik mittags erreicht wird, während beim Nachlassen der Insolation am Abend die Ozonzerstörung wieder die Oberhand gewinnt. Ähnliche Kurven liegen auch in Weißenau in ländlicher Lage vor.

In Abb. 6 sieht man dasselbe. Die abendlichen Anstiege am 19. und 20. 7. in Staufen und in Tübingen sind durch die Turbulenz bei Gewittern bedingt.

Abb. 7 zeigt, daß im Winter bei bedecktem Himmel und Windstille in Tübingen zuweilen überhaupt nichts mehr von dem damals auch in Arosa geringen Ozon

anzutreffen ist. Die Heizgase der Stadt liegen bei solchem Wetter als sichtbarer Dunst über der Stadt und zerstören das Ozon sehr wirksam.

Schließlich zeigt Abb. 8 noch einen der seltenen Fälle, in welchem Arosa weniger Ozon hat als Tübingen und der Baldenweger Hof bei Freiburg i. Br. Hier ist am

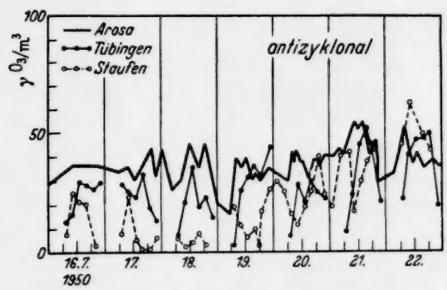


Abb. 6.

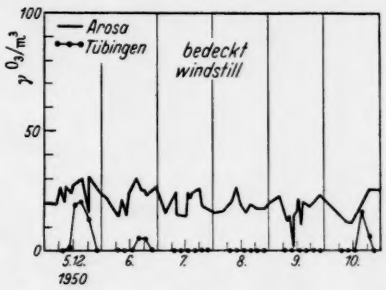


Abb. 7.

17. 1. 1951 unter einer starken Inversion (ab 2000 m) ein neuer ozonreicherer Luftkörper vorgedrungen, während Arosa und St. Blasien sich noch in dem alten ozonarmen Luftkörper befinden.

Diese Ergebnisse bestätigen alle deutlichst die Auffassung, daß das Ozon am Boden rasch zerstört wird [1], [2] und zeigen erneut [3], [5], daß das Ozon als ziemlich beständiges Kennzeichen eines Luftkörpers wichtige meteorologische Aufschlüsse geben kann. Allerdings sollte man nicht am Boden und in Städten, sondern in der freien Atmosphäre messen. Die früher von uns durchgeführten Messungen im Flugzeug hatten dafür trotz ihrer geringen Zahl bereits eindeutige Belege ergeben [4], [5].

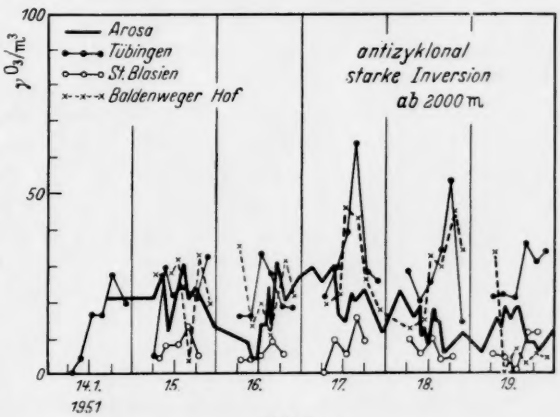


Abb. 8.

Die von uns entwickelte Meßmethodik wurde bereits veröffentlicht [6]. In den 2 Jahren seit dem Abschluß dieser Veröffentlichung wurden noch weitere Vereinfachungen erzielt, welche in diesem Zusammenhang eine kurze Darlegung angebracht erscheinen lassen.

Wie viele andere Autoren benutzen wir die Oxydation von Jodkalium durch das Ozon in neutraler Lösung

$$O_3 + 2\,KJ + H_2O \rightarrow O_2 + 2\,KOH + J_2.$$

Die Luft wird in feinen Bläschen durch 3 cm³ einer 2%igen KJ-Lösung gesaugt. Dieser Lösung ist außerdem noch $Na_2S_2O_3$ zugesetzt, welches sofort mit dem durch Oxydation entstehenden Jod reagiert:

$$2\,Na_2S_2O_3 + J_2 \rightarrow 2\,NaJ + Na_2S_4O_6.$$

VOL. 2 1952

Durch die Oxydation des KJ durch das Ozon entsteht also ein Verlust an Natriumthiosulfat. Diesen messen wir nachher und können daraus auf die Menge des eingesaugten Ozons schließen. Der Umweg über das Natriumthiosulfat ist notwendig, weil freies Jod J_2 durch den Luftstrom aus der Lösung entfernt würde. Natriumthiosulfat unterliegt diesem Verlust nicht.

Abb. 9 zeigt das für diesen Zweck entwickelte Reaktionsgefäß. Aus einer Vorratslösung 2%iger KJ-Lösung mit einem Zusatz von Natriumthiosulfat (ungefähr 1 cm³ $n/100$ zu 750 cm³ KJ-Lösung) werden mit der Pipette genau 3 cm³ in das Gläschen (*4*) gegeben und dieses an den Glasteil (*2*) mit dem Glasschliff angesetzt. Wird nunmehr bei (*3*) Luft abgesaugt, so tritt die zu untersuchende Luft bei *E* ein, strömt im Innern von Teil (*2*) nach unten und gelangt durch die Glasfritte fein verteilt in die Lösung. Das Gefäß *2* ist so ausgebildet, daß selbst bei Pumpgeschwindigkeiten bis zu 2 l pro min noch keine Lösung an den Schliff (*3*) verspritzt wird und alles Ozon restlos zur Reaktion kommt. Dies wurde mit hintereinander geschalteten Reaktionsgefäßen eingehend geprüft. Erst bei Lufttemperaturen unter 3° C beginnt die Reaktion unvollständig zu werden. Deshalb wird im Winter die Luft über ein 1 m langes Glasrohr angesaugt, auf dessen Außenseite eine Heizwicklung mit 20 w und eine Wärme-Isolierung aufgebracht sind.

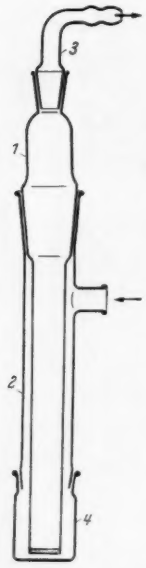


Abb. 9.

Nach Beendigung der Reaktion wird durch Einblasen beim Schliff (*3*) zunächst die Lösung in das abgehobene und etwas tiefer gehaltene Gläschen (*4*) zurückgedrückt und durch Einspritzen von etwa 3 cm³ bidestillierten Wassers durch den oberen Schliff die hängen gebliebene Lösung ebenfalls in das Gläschen (*4*) gespült.

Nunmehr ist der Natriumthiosulfatgehalt dieser Lösung zu bestimmen und mit demjenigen von 3 cm³ der Vorratslösung zu vergleichen. Die Differenz ist das Äquivalent des in der Lösung durch das eingesaugte Ozon gebildeten Jods.

Ist diese Differenz J γ Jod äquivalent und wurden N Liter Luft angesaugt, so beträgt der Ozongehalt der Luft

$$C_{O_3} = \frac{J\ [\gamma\ \text{Jod}] \cdot 1000\ [\text{Liter/m}^3]}{N\ [\text{Liter}] \cdot 5{,}23 \left[\frac{\gamma\ \text{Jod}}{\gamma\ O_3}\right]} J = \cdot \frac{1{,}93}{N} \left[\frac{\gamma\ O_3}{\text{m}^3}\right].$$

Der entscheidende Fortschritt wurde nun durch die Entwicklung eines Jodmeters erzielt, mit welchem Jod in KJ-Lösung und ebenso Natriumthiosulfat in KJ-Lösung mit sehr großer Genauigkeit und Empfindlichkeit gemessen werden können. Dies sei hier kurz beschrieben [7].

Taucht man zwei Platinelektroden in eine Jodkaliumlösung und legt eine Spannung an, die kleiner als die Zersetzungsspannung ist, so bilden sich sog. reversible Jodelektroden mit einer Polarisationsspannung gleich der angelegten Spannung. Der Strom geht auf 0 zurück. Wenn nun aber freies Jod (praktisch als J_3) in der Lösung ist, so erhält man einen Strom, welcher bei rascher Bewegung der Lösung gegen die Elektroden in einem sehr weiten Bereich (sechs Größenordnungen) der Konzentration des freien Jods proportional ist. Die Lösung wird durch diesen Meßvorgang nicht verändert. Bei der praktischen Ausführung dreht ein Grammophon-

motor mit Drehzahlregelung einen Halter, in welchem die Gläschen so wie sie vom Reaktionsgefäß kommen eingesetzt werden können. Wir tauchen die Elektroden in die sich drehenden Gläschen nahe der bewegten Glaswand ein und legen eine Spannung von 0,18 v an. Dabei entsteht zunächst ein Polarisationsstrom, der jedoch rasch auf sehr kleine Werte absinkt, wenn in der Lösung Natriumthiosulfat im Überschuß ist.

Gibt man jetzt Jod zu, so ändert sich nichts, bis alles Natriumthiosulfat aufgebraucht ist, welches ja das eingebrachte Jod sofort in Natriumtetrathionat umsetzt. Die Jodmenge, welche über das Äquivalent der ursprünglichen Natriumthiosulfatmenge hinaus zugeführt wird, bewirkt einen Anstieg der Jodkonzentration proportional zur überschüssigen Menge und wir erhalten einen ihr proportionalen Strom.

Das freie Jod geben wir nicht chemisch zu, sondern erzeugen es in der Lösung durch Elektrolyse über ein zweites Elektrodenpaar.

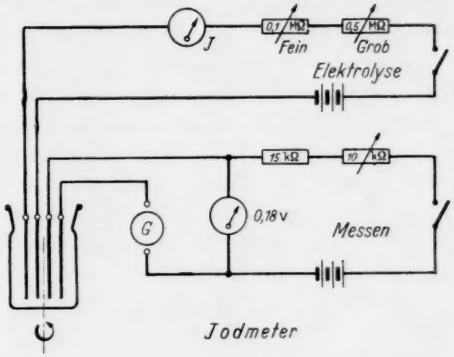


Abb. 10.

Abb. 10 zeigt die Schaltung der Anordnung. Der Elektrolysenstrom kann am Instrument *J* abgelesen und durch die Widerstände grob und fein geregelt werden. Bei 38 μA wird nach dem FARADAYschen Gesetz in 20 sec 1 γ Jod abgeschieden.

Abb. 11 zeigt nun den Meßstrom über die Elektroden 1 und 2 in Abhängigkeit von der Zeit nach dem Einschalten bei verschiedenen Elektrolysenströmen, aber jeweils bei derselben Menge Vorratslösung, also auch bei demselben Natriumthiosulfatgehalt. Man erkennt, daß der Anstieg des Stromes zwar nicht scharf einsetzt, aber dann ausgezeichnet linear wird, so daß der Schnittpunkt dieser Anstiegsgeraden mit dem Nullstrom sehr gut definiert ist. Das ist neben der erhöhten Empfindlichkeit der entscheidende Fortschritt gegenüber der Titration mit Farbumschlag. Der in Abb. 11 durch Schraffur hervorgehobene Übergang der Kurven kommt dadurch zustande, daß die Lösung in der Praxis nicht in jedem Augenblick ideal durchmischt werden kann. Das Jod entsteht an der Anode der Elektrolyse und muß erst mit der Lösung gemischt werden. Je kleiner der Elektrolysenstrom, desto schärfer wird der Übergang, aber desto länger dauert die Messung.

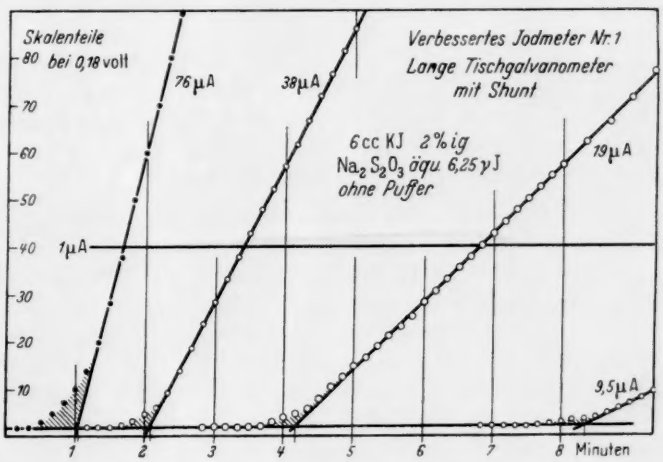


Abb. 11.

Die Umrechnung der verschiedenen Fußpunkte ergibt jeweils dieselbe Jodmenge. Es sind hier Eichmessungen eines bestimmten Gerätes aufgezeichnet, das für rasche Messungen mit einer Genauigkeit von $^1/_{10}\,\gamma$ Jod bestimmt war. Es kann auch für eine Genauigkeit von $^1/_{1000}\,\gamma$ Jod eingerichtet werden*.

Eine Verdünnung der Lösung mit Wasser oder eine andere Temperatur oder ein langsameres Rühren oder ein Strommesser *G* in Abb. 10 mit anderer Eichkonstante bewirken nur eine Änderung der Steilheit der Anstiegsgeraden im Strom-Zeit-Diagramm. Der das Resultat bestimmende Fußpunkt wird aber dadurch nicht beeinflußt. Dies bedeutet eine große Vereinfachung der Messungen. Jede Einzelmessung ist in 3 bis 4 min erledigt und kann durch ein Registrierinstrument bei automatischem Ablauf dokumentarisch festgehalten werden. Das Spülen der Elektroden und die Auswertung benötigen keine halbe Minute, wenn man den Elektrolysenstrom so wählt, daß ein einfaches Zahlenverhältnis zwischen den Einheiten des Zeitmaßstabes und der gesuchten Größe besteht.

Die Empfindlichkeit des Jodmeters erlaubt uns die Ausführung der Ozonmessungen mit weniger als 1 l Luft. Aber dieses Luftvolumen muß genau definiert sein. Deshalb benutzen wir doch etwa 10 l Luft, die mit einer Membranpumpe angesaugt werden. Diese stellt nach einer fest eingestellten Umdrehungszahl automatisch ab. Außerdem betätigt eine neuere Ausführung noch im Augenblick des Abschaltens einen Umschalthahn, welcher die Pumpe dabei auf den Ausgang eines anderen Waschrohres umschaltet. Wir benutzen z. Z. in Weißenau und in Tübingen je 12 Waschrohre und setzen mit einer Kontaktuhr alle 2 Stunden eine Messung in Gang, die jeweils 5 min dauert.

Man muß dann nur 1mal am Tag die 12 Lösungen auswerten und durch neue ersetzen. Dies macht eine Person bequem in 1 Stunde. Wenn die Lösungen vor Licht geschützt bleiben, entstehen in dieser Zeit keine Fehler. Um so mehr, als jeweils bei den Waschrohren ein weiteres Gläschen mit gleicher Lösung aufbewahrt wird und bei der Auswertung zur Kontrolle dient.

Die Vermeidung von Normallösungen ist für die Praxis ein großer Vorteil. Die Zuverlässigkeit des Meßverfahrens ergibt sich aus der Tatsache, daß beim Fehlen von Ozon genau der Wert 0 herauskommt und daß *genau* gleichzeitig am selben Ort vorgenommene Messungen nur höchstens um 5% verschieden ausfallen. Diese Streuung geht hauptsächlich auf die Streuung des gepumpten Luftvolumens zurück. Die größte Unsicherheit bedingt in der Praxis tatsächlich die Umrechnung des Pumpenvolumens auf Luftvolumen, besonders wenn die Luft angewärmt werden muß. Angesichts der Schwankung des Ozongehaltes sind diese kleinen meßtechnischen Unsicherheiten belanglos.

Störungen treten in unserem Gebiet nur auf, wenn Rauchgase oder Industriegase an den Meßort gelangen. Man muß dann die Lösung puffern [1 cm³ $n/10$ (Na_2HPO_4 + NaH_2PO_4)]. An der See ist noch ein Einfluß von Cl_2 zu erwarten, der durch besondere Maßnahmen abzutrennen ist. Der Puffer beeinträchtigt die Ozonmessung und die Elektrolyse des KJ nicht. Stickoxyde oxydieren das neutrale KJ nach unseren Laboratoriumsversuchen nicht. In der Natur konnten wir sie nicht in meßbarer Menge nachweisen.

* Der Aufwand an Instrumenten und Zeit für eine Messung wird aber dann sehr viel größer.

Für sehr genaue absolute Messungen von Ozonmengen sind noch andere Reaktionsmöglichkeiten zwischen KJ und O_3 zu berücksichtigen. Untersuchungen darüber sind im Gange. Ihr Einfluß scheint aber bei neutraler Lösung klein zu sein. Es ist ein Vorteil unserer Methode, daß bei der Reaktion nur sehr wenig von der zugegebenen KJ-Menge verbraucht wird und deshalb die Wasserstoffionenkonzentration praktisch konstant bleibt.

Zusammenfassung

Mit einer leicht zu handhabenden chemischen Meßmethode wurde an mehreren Stationen seit längerer Zeit regelmäßig die lokale Ozonkonzentration der atmosphärischen Luft gemessen. Es werden einige typische Beispiele gezeigt, aus welchen die Ozonzerstörung in der Bodenschicht und der beherrschende Einfluß des Luftaustausches mit höher gelegenen Schichten anschaulich zu erkennen ist. Die Meßmethode wird kurz beschrieben.

References

[1] Auer, R.; Gerlands Beiträge zur Geophysik 1939 **54** 137. [2] Regener, E.; Forschungs- und Erfahrungsberichte des Reichswetterdienstes A (Nr. 9), 1941; Meteorolog. Z. 1943 **60** 253. [3] Ehmert, A. u. H.; Forschungs- und Erfahrungsberichte des Reichswetterdienstes A (Nr. 13) 1941. [4] Ehmert, A.; Forschungs- und Erfahrungsberichte des Reichswetterdienstes A (Nr. 13) 1941. [5] Ehmert, A.; Sitzungsbericht der Sondertagung "Ozon" des Reichswetterdienstes am 17. und 18. April 1944. [6] Ehmert, A.; Meteorologische Rundschau 1951 **4** 64–68. [7] Vgl. Ehmert, A.; Z. Naturforschung 1949 **4b** 321–327.

Note: [2], [3], [4] und [5] im Neudruck in Ber. Dtsch. Wetterdienst. US-Zone (Nr. 11) 1949.

VOL. 2 1952

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 196 to 199. Pergamon Press Ltd., London

Indices of geomagnetic activity

of the Observatories ABINGER (*Ab*), ESKDALEMUIR (*Es*) and LERWICK (*Le*)

August to December 1951

The figures given on pages 196 to 199 represent the K-indices for three-hour intervals, beginning with 00.03 hrs for the first and ending with 21–24 hrs for the eighth figure.

August 1951

Day	*Ab* Range for $K = 9$: 500 γ		Sum	*Es* Range for $K = 9$: 750 γ		Sum	*Le* Range for $K = 9$: 1000 γ		Sum	Day
	K-Indices			*K-Indices*			*K-Indices*			
1	4223	3666	32	4213	3556	29	4312	3446	27	1
2	5432	3333	26	5332	3333	25	5532	3223	25	2
3	2123	3222	17	2123	3222	17	2212	2112	13	3
4	3322	4344	25	3312	3333	21	3422	3334	24	4
5	4223	3432	23	4212	3422	20	5312	2422	21	5
6	3133	3333	22	3021	3313	16	3122	3332	19	6
7	1323	2333	20	1312	2323	17	1312	2324	18	7
8	2223	2223	18	2222	2122	15	2222	2112	14	8
9	2223	3333	21	2112	3323	17	2112	3323	17	9
10	2222	3323	19	1122	3223	16	1122	3223	16	10
11	2233	4333	23	2222	3333	20	2222	3334	21	11
12	3333	3334	25	3232	4334	24	4332	3343	25	12
13	2444	4451	28	2443	4451	27	2543	5562	32	13
14	2113	2334	19	1113	3333	18	1113	3324	18	14
15	3333	3453	27	3232	3452	24	4331	3343	24	15
16	1355	4423	27	1255	3323	24	1244	3322	21	16
17	3233	3523	24	3121	3533	21	2222	3434	22	17
18	2222	2211	14	2221	2101	11	2221	1211	12	18
19	0133	3432	19	0132	3322	16	1132	3332	18	19
20	5543	3354	32	5544	3354	33	6553	3454	35	20
21	5444	4545	35	5334	4655	35	7533	5766	42	21
22	5443	3343	29	5432	4342	27	7542	4443	33	22
23	3123	3535	25	3123	3535	25	4123	3435	25	23
24	4334	3344	28	4333	3234	25	5533	3233	27	24
25	4454	3552	32	3444	2452	28	4544	3452	31	25
26	4344	4445	32	4343	4444	30	5344	3344	30	26
27	2333	3455	28	2333	3454	27	3222	3354	24	27
28	3334	4344	28	2223	3344	23	2222	3234	20	28
29	3333	3423	24	3223	3323	21	4222	3323	21	29
30	1113	3333	18	0002	2333	13	1112	2333	16	30
31	2234	3334	24	2123	4324	21	2123	4423	21	31

September 1951

Day	*Ab* Range for $K = 9$: 500 γ		Sum	*Es* Range for $K = 9$: 750 γ		Sum	*Le* Range for $K = 9$: 1000 γ		Sum	Day
	K-Indices			*K-Indices*			*K-Indices*			
1	4323	3313	22	2123	2213	16	2112	2213	14	1
2	2223	2222	17	2222	2212	15	2112	2212	13	2
3	3233	3113	19	3122	3113	16	3122	2113	15	3
4	2333	3323	22	2222	3222	17	2222	3213	17	4
5	3133	3244	23	2122	2244	19	2122	2234	18	5
6	3243	3343	25	4133	3243	23	4232	3233	22	6
7	3123	3342	21	2012	2342	16	3112	2232	16	7
8	1133	3433	21	1122	3332	17	1221	3333	18	8
9	3223	4433	24	3223	3323	21	3222	4344	24	9
10	4444	4332	28	4343	4331	25	4434	5341	28	10
11	3333	3446	29	3232	3446	27	4232	4467	32	11
12	4333	4435	29	4333	4435	29	6433	4437	34	12
13	5433	4445	32	5432	3336	29	8542	4458	40	13
14	4443	4333	28	4433	4232	25	5533	3233	27	14
15	4353	4434	30	4343	4423	27	4343	4433	28	15
16	3545	5556	38	3545	5656	39	5555	6767	46	16
17	5344	4435	32	4333	4435	29	5434	4434	31	17
18	4333	3553	29	3113	2553	23	5432	2554	30	18
19	3233	6754	33	3133	5774	33	3223	4885	35	19
20	4465	5665	41	4455	5666	41	5555	5568	44	20
21	4556	5555	40	5556	5556	42	8556	5556	45	21
22	5555	5465	40	4545	5465	38	6654	5557	43	22
23	4544	4455	35	5543	3454	33	7643	4455	38	23
24	5544	4344	33	4444	4343	30	4443	4343	29	24
25	3344	6678	41	3334	6889	44	3334	8889	46	25
26	7644	3423	33	9633	3322	31	9733	3323	33	26
27	5555	4333	33	5545	4333	32	5744	4324	33	27
28	4223	2102	16	4212	1102	13	6311	2212	18	28
29	3222	4543	25	3121	3543	22	3121	3743	24	29
30	3322	2111	15	3222	2101	13	4222	2111	15	30

October 1951

Day	K-Indices		Sum	K-Indices		Sum	K-Indices		Sum	Day
1	2212	1322	15	2011	1222	11	1111	1322	12	1
2	3333	2324	23	3323	2214	20	3332	2213	19	2
3	2212	3133	17	2112	2132	14	1102	2132	12	3
4	1123	1232	15	0012	1122	9	1001	1122	8	4
5	3011	1122	11	3011	1122	11	3011	1122	11	5
6	2112	1100	8	2011	0000	4	2101	0001	5	6
7	1123	3565	26	1112	4465	24	1222	3677	30	7
8	6554	5444	37	5454	4444	34	6645	4445	38	8
9	4334	5335	30	3233	4335	26	4423	3335	27	9
10	4335	3555	33	4334	3545	31	5533	3554	33	10
11	4234	3330	22	3233	3320	19	5332	2341	23	11
12	1224	4332	21	1213	4322	18	1223	3232	18	12
13	3334	3345	28	2323	2245	23	2323	2246	24	13
14	5233	1354	26	5223	1354	25	5222	1364	25	14
15	3222	2144	20	2222	1134	17	3221	1134	17	15
16	3333	3345	27	2323	2334	22	2223	2345	23	16

VOL. 2 1952

October 1951 (Continued)

Day	*Ab* Range for $K = 9$: 500 γ		Sum	*Es* Range for $K = 9$: 750 γ		Sum	*Le* Range for $K = 9$: 1000 γ		Sum	Day
	K-Indices			*K-Indices*			*K-Indices*			
17	5555	5675	43	4544	4574	37	7554	4775	44	17
18	4454	4654	36	3444	4544	32	5544	4544	35	18
19	5444	4565	37	5434	4465	35	5543	3476	37	19
20	5233	3352	26	5233	3252	25	7323	3242	26	20
21	3334	2232	22	3333	2232	21	2323	2233	20	21
22	3333	3544	28	3222	3543	24	3322	4543	26	22
23	3223	3333	22	3112	2323	17	3222	2323	19	23
24	2132	1222	15	2101	1112	9	1111	1112	9	24
25	2123	2211	14	2001	1101	6	2111	1110	8	25
26	1133	3235	21	0023	2125	15	0122	2226	17	26
27	3323	3101	16	3312	3001	13	3322	3011	15	27
28	0335	6885	38	0224	5894	34	0234	8995	40	28
29	3333	3211	19	3332	3101	16	5222	3211	18	29
30	1233	2210	14	1223	2110	12	1232	2110	12	30
31	0122	1112	10	0001	1101	4	0111	1111	7	31

November 1951

Day	*K-Indices*		Sum	*K-Indices*		Sum	*K-Indices*		Sum	Day
1	1022	1233	14	1022	2113	12	1012	1112	9	1
2	3333	3333	24	2323	2333	21	2322	2324	20	2
3	3434	4246	30	3433	3345	28	3433	3346	29	3
4	4543	4434	31	4432	3434	27	5532	3335	29	4
5	3323	2234	22	3323	1234	21	3322	1233	19	5
6	3333	3364	28	2232	3363	24	2222	3565	27	6
7	4444	2333	27	4432	2322	22	5532	2323	25	7
8	3323	2232	20	3211	1022	12	4211	1032	14	8
9	1333	3433	23	1233	3332	20	1223	3432	20	9
10	2122	2111	12	2001	1000	4	2101	1100	6	10
11	0011	3134	13	0010	2133	10	0000	2123	8	11
12	4343	3523	27	4242	3412	22	5332	4522	26	12
13	3245	4555	33	2234	3565	30	2233	3676	32	13
14	4343	4565	34	5343	3465	33	7332	3487	37	14
15	4434	3433	28	4323	3433	25	4333	3434	27	15
16	3333	2331	21	3222	2321	17	4222	2332	20	16
17	1322	3365	25	1122	2264	20	1222	2264	21	17
18	3112	1132	14	2012	1122	11	2112	1133	14	18
19	1212	2223	15	1101	1223	11	1111	1213	11	19
20	4123	2344	23	3211	1244	18	3211	1345	20	20
21	2322	1134	18	2211	1033	13	2211	1134	15	21
22	4133	2344	24	3112	2234	18	3111	1243	16	22
23	3222	3454	25	3212	2343	20	3211	2344	20	23
24	3312	3432	21	3210	2431	16	2211	3422	17	24
25	2133	3435	24	2122	2325	19	3132	3335	23	25
26	2122	4242	19	2112	3242	17	2112	3242	17	26
27	3222	1123	16	2121	0113	11	2111	1113	11	27
28	2232	4344	24	2222	3234	20	3222	3244	22	28
29	3234	4444	28	3223	3444	25	3323	3444	26	29
30	4243	4332	25	4122	3221	17	4122	3332	20	30

December 1951

Day	*Ab* Range for $K = 9$: 500 γ		Sum	*Es* Range for $K = 9$: 750 γ		Sum	*Le* Range for $K = 9$: 1000 γ		Sum	Day
	K-Indices			*K-Indices*			*K-Indices*			
1	3223	3333	22	2213	3233	19	2212	2233	17	1
2	3332	2333	22	3332	2332	21	3232	2443	23	2
3	4212	2234	20	3211	1223	15	4311	2324	20	3
4	3333	4455	30	2223	3455	26	2222	3445	24	4
5	3313	3323	21	3312	2323	19	3212	2323	18	5
6	1112	1131	11	1000	0031	5	1001	1132	9	6
7	2123	3444	23	2102	2343	17	1102	2443	17	7
8	6345	4565	38	5345	4464	35	6334	4475	36	8
9	3434	5553	32	3434	4553	31	4434	4653	33	9
10	1344	3545	29	2244	3544	28	2333	3545	28	10
11	4344	4435	31	4344	3434	29	4244	3434	28	11
12	2232	3133	19	2122	3123	16	2122	3123	16	12
13	2322	2122	16	1221	2012	11	1222	1112	12	13
14	1312	1144	17	2201	1034	13	1111	0035	12	14
15	4233	3353	26	4233	3352	25	4323	3243	24	15
16	3332	2233	21	2322	2223	18	2321	2213	16	16
17	2322	3444	24	1312	3444	22	2312	2554	24	17
18	4332	2445	27	3322	2434	23	3422	2434	24	18
19	3234	4444	28	3223	4434	25	3223	3434	24	19
20	4334	2132	22	3333	2132	20	3323	1232	19	20
21	2122	1224	16	2100	1114	10	2100	1114	10	21
22	3423	3555	30	2413	3554	27	3412	3765	31	22
23	4324	1231	20	4223	2220	17	5312	2211	17	23
24	1122	1222	13	1001	0011	4	2011	0021	7	24
25	1112	1112	10	1001	1102	6	1101	1112	8	25
26	0011	0111	5	1001	0000	2	0011	1111	6	26
27	0313	3125	18	0202	3014	12	1212	2113	13	27
28	4664	5454	38	3654	4444	34	3665	5454	38	28
29	3212	4321	18	3101	3221	13	4101	3221	14	29
30	0112	2224	14	0002	2213	10	0012	1113	9	30
31	3245	4565	34	3234	3455	29	3234	3565	31	31

VOL. 2 1952

Book review

HARANG, L.; **The Aurorae.** Vol. I of the International Astrophysics Series. General Editors: M. A. ELLISON, Edinburgh, and A. C. B. LOVELL, Manchester. Chapman and Hall, 1951. 25/- net.

Dr. LEIF HARANG, a member of the well-known school of Norwegian geophysicists, is well qualified to write on The Aurorae in the first volume of the new *International Astrophysics Series.* He was, for many years, Director of the Auroral Observatory at Tromsø, in Northern Norway, and has himself conducted an extensive series of researches on the visual and spectroscopic manifestations of aurorae as well as on their accompanying magnetic and ionospheric perturbations. His book, which is short and concise, deals first with methods of locating aurorae in space and with the identification of the spectrum of the aurora and the night sky—investigations in which his colleagues STØRMER and VEGARD have, respectively, played such outstanding rôles. Then there follow chapters on the terrestrial magnetic storms which usually accompany auroral displays and on the solar corpuscular theory of the origin of both. Here the theoretical work of BIRKELAND and STØRMER on the influence of the earth's magnetic field on the trajectory of an electrically charged particle is given due treatment; though the later work of CHAPMAN and FERRARO on the same subject, in which a *neutral* stream of solar particles is envisaged, is not given comparable treatment.

The two final chapters deal with the physics of the upper atmosphere as disclosed by all types of geophysical observations. In one of them the author describes some elegant experiments of his own

on the distribution of auroral luminosity with height, from which he has been able to estimate the value of the scale height of the atmosphere at different levels, and from which, further, he derives values of the density and pressure of the air at all levels between 75 and 800 km above the earth's surface.

Dr. HARANG's volume on the aurorae can be recommended as a reliable and friendly guide to the subject. It is hoped, however, that a second edition will soon be called for, so that results appearing more recently than December 1949, the date when the present text was finished, can be included in it.

E. V. A.

CORRECTIONS

1. *In the paper by* D. H. SHINN *and* H. A. WHALE *published in the present volume of this Journal on pages 85 to 105, the following items require correction:*

Page 92. The third line of equation (7) should read as follows:

$$\text{Total field} = H \propto \sqrt{(1 + 3\sin^2\lambda)} \propto \sqrt{1/(4 - 3\sin^2\varphi)} \qquad (7)$$

Page 103. The heading to Table 5 should read as follows:

Table 5. Increase in apparent height, $h' - h_0$, due to the presence of a magnetic field. Ordinary ray. Linear layer. In England

2. *In the paper by* F. H. NORTHOVER, *published in the present volume of this Journal on pages 106 to 129, the following items require correction:*

Page 106. Line 4 under INTRODUCTION, high "level inversions", should read: "high level inversions".

Page 114. The latter half of the first line immediately following equation (46) should read as follows:

$$\text{and similarly } v_{-s-\frac{1}{2}}(\xi) = e^{s\pi i} v_{s-\frac{1}{2}}(\xi). \text{ Thus}$$

Page 115. The first expression in the footnote should read

$$\zeta_\eta' - \zeta'_\eta$$

Page 118. Equation (67): -1 at the end of the first line should read $-i$, and the small italic i in the equation next following should be deleted.

Page 119. Equation (80): the first symbol should read ζ not S. Equation (84): there should be a closing bracket at the end of the equation.

Page 126. Under APPENDIX II(A), in the first line $\varkappa_1$ should be replaced by x_1 in both places where it occurs.

The heading APPENDIX III(B) should read: APPENDIX II(B).

Page 127. The line immediately above "*Case 2*" should read as follows:

$$\text{since } \zeta_\nu' - \zeta'_\nu = i$$

Page 128. "$\tau < 22\lambda$ metres", should read: "$\tau > 22\lambda$ metres".

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 201 to 215. Pergamon Press Ltd., London

Electric conductivity and small ion concentration of the atmosphere at one metre above ground and conductivity at ground level

GEORGE A. O'DONNELL

Iona College, New Rochelle, N.Y.

(Received 10 October 1951)

ABSTRACT

The diurnal variation of the polar conductivities of the atmosphere was obtained at ground level and at 1 m above ground at five locations in and within 120 miles of New York City. The diurnal variation in the concentration of small ions of both signs at 1 m above ground was found simultaneously with the conductivity measurements. Results show that the polar conductivities at ground level are not equal to the respective values at 1 m and the ratios of the polar ground conductivities to the corresponding metre values are not constant. Factors such as meteorological changes and local air conditions do not affect conductivity at both levels equally. The total conductivity at ground level is in general not equal to the total conductivity at 1 m nor is the ratios of these quantities constant. Some of the factors causing variation in the concentration of small ions have a parallel effect on the average mobilities* of the ions in a given mobility range. Thus variation in conductivity is dependent not only on the concentration of small ions but also on their average mobilities. An explanation is offered for the apparent fact that while the electric field of the earth shows little difference from the ground level up 1 m above ground, conductivity may differ considerably between the two levels.

VOL. 2 1952

INTRODUCTION

Measurements of atmospheric conductivity at ground level and at 1 m above ground have been made at Kew by WATSON [1], SCRASE [2] and HOGG [3] and at Glencree by P. J. NOLAN [4]. WATSON used two WILSON Universal Electrometers one flush with the ground and the other at about 1 m (130 cm). He found that conductivity at ground level to be on the average about 25% greater than at 1 m. SCRASE used the same method with improved facilities and obtained approximately the same results. He believed the conductivities should be equal at the two levels and attributed the lower 1-m value to defects in experimental procedure. These investigators measured the positive conductivity only. HOGG obtained the polar conductivities at 1 m by the aspiration method and found their sum to be about equal to the positive conductivity at ground level measured with the WILSON apparatus. NOLAN found the total conductivity at 1 m by the aspiration method to exceed that measured at the ground by the WILSON method by about 12%. To account for this difference he postulated an upward convection current due to turbulence the magnitude of which, on the average, was 12% of the air-earth current.

The results of many investigations have shown that the concentration of small ions and the electric conductivity of the atmosphere vary considerably over land from place to place in the lowest levels of the earth's atmosphere. That this variation depends upon such local factors as the amount of dust, smoke, water-vapour and condensation-nuclei in the free air has been fairly well established ([5], [6], [7]). However, simultaneous measurements of the concentration of small ions and of the polar conductivities have not been numerous. If the field of the earth remains approximately constant for a period of observation then variations in conductivity are caused by variations in the small ion content of the atmosphere and in the average mobilities of these ions.

* *Average mobility* is the mobility calculated from one observation of conductivity and ion concentration taken simultaneously.

The purpose of this investigation was first, to obtain the diurnal variation of the polar and total conductivities at ground level and at 1 m above ground and second, to make simultaneous measurements of the small ion concentration at 1 m and compare these elements by means of the relation $\lambda_{\pm} = n_{\pm} e u_{\pm}$.

For the above purposes, observations of conductivity and concentration of small ions were made at five locations. A brief description of each will be given under "Results of Observations." By choosing locations in and at near and distant points from New York City, it appeared possible to obtain a kind of gradation of air conditions as they varied from city air to purer country air. Further, it seemed that this procedure would show some relationships between the ground and 1-m conductivities which were not influenced by one set of local conditions. At the various locations observations were made on extensive lawns except at West Park where the point of observation was the centre of an eight-acre tract of grazing land. The apparatus was always set up under the branches of large trees as a protection from the field of the earth.

Two aspiration instruments with the collecting cylinders horizontal, were used for ion counting and conductivity measurements. The conductivity apparatus was of the Gerdien type. A right-angled elbow of galvanized iron tubing was made to fit tightly over the end of the outer cylinder. The long arm was 1 m in length measured from the axis of the cylinder to the open end. This elbow could be turned about the axis of the cylinder. Preliminary experiments with and without the elbow, showed that the rate of air flow was sufficient to prevent loss of ions by diffusion to the walls. Although for ground measurements the end of the elbow was usually 6 cm from the ground, many observations up to 15 cm did not change the results. Evidently the method of drawing air directly up insured its coming from a level in contact with the ground. For observations at the 1 m level, the elbow was turned to a horizontal position with the intake 1 m above the ground. Thus, measurements at both levels were under identical experimental conditions and if differences in their values appear in the results it is more certain they are real and not apparent.

Usually, alternate observations of conductivity of one sign at ground level and at 1 m with the corresponding ion concentration were made for 24 hrs and then followed 24 hrs of observation of the other sign. At times the positive and negative elements were measured alternately, but for low values of conductivity the time to complete a cycle was too long. We hoped that many measurements though taken on different days would given mean values sufficiently representative for each location. Conductivities are expressed in electrostatic units with the factor 10^{-4} omitted from the tables.

The following symbols are used:

n_1 = positive small ions per cm^3
n_2 = negative small ions per cm^3
λ_1 = positive conductivity at 1 m
λ_{1g} = positive conductivity at ground level
λ_2 = negative conductivity at 1 m
λ_{2g} = negative conductivity at ground level
$\Lambda = \lambda_1 + \lambda_2$ = total conductivity at 1 m
$\Lambda_g = \lambda_{1g} + \lambda_{2g}$ = total conductivity at ground level
u_1 = average mobility of positive small ions
u_2 = average mobility of negative small ions

Results of Observations

1. Iona

Iona College, in the City of New Rochelle, 20 miles north-east of New York City centre, is situated in a residential district of individual homes. At this location observations were made during June and July (1950) but not for continuous 24-hr periods as only one observer was available. Nevertheless, the 24 hrs were covered for all elements except the negative conductivity at ground level but enough measurements of this element were made during the day and night to establish a relationship to λ_2 for these periods (Table 1). Here many observations alternating between the positive quantities λ_1 and n_1, and the negative λ_2 and n_2 were also made.

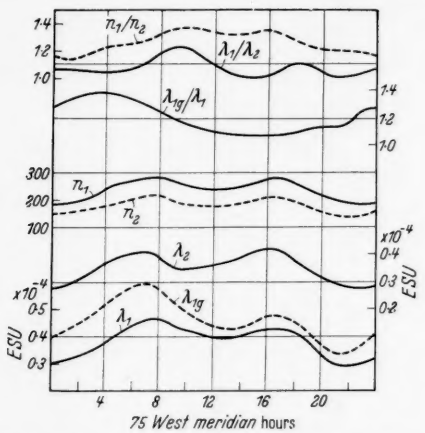


Fig. 1. Atmospheric-Electric results from Iona.

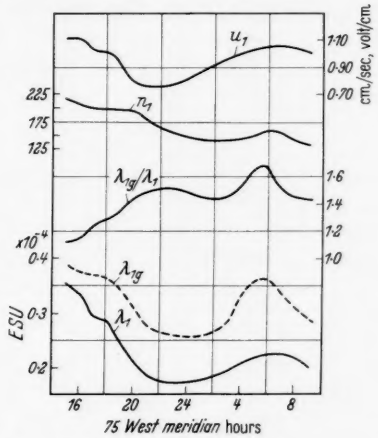


Fig. 2. Variation of positive elements under stable conditions.

Results are given in Fig. 1. The morning maximum conductivity occurs between 06 h and 08 h and the minimum between 22 h and 24 h local time. During the night at this time of year the wind is seldom greater than 1 (Beaufort scale) but in the daytime it is gentle to moderate (Beaufort 2–4) being strongest in the late afternoon. This would account for the rise in conductivity and ion concentration between 15 h and 16 h. The sudden drop after 17 h is assumed due to domestic fires in the neighbourhood at this time. The concentration of positive ions and the positive conductivities at both levels were always greater than the corresponding negative values. The positive ground conductivity was always greater than the positive 1-m conductivity while negative ground conductivity was less than that at 1 m. The total conductivity at both levels was greater during the day than at night but all the ratios of the polar and total conductivities at ground level to the corresponding values at 1 m were greater for the night period (Table 1). The ratio, λ_{1g}/λ_1, increases when λ_1 decreases and for values of conductivity at and near the time of maximum (Fig. 11). This would mean that λ_1 is more variable than λ_{1g} except at the time of maximum when λ_{1g} becomes more variable.

The diurnal variation in the concentration of small ions as expected shows the same variation as the values of the corresponding conductivities but the morning maximum occurs about an hour later than the maximum for conductivity while

VOL. 2 1952

the time of minimum ion concentration is the same as that for conductivity. Day and night means are given in Table 2 with mean mobilities. The minimum average mobility for positive ion at this location was found to be 1·06 cm/sec,v/cm and the maximum was 1·19. The respective values for negative ions were 1·22 and 1·48.

In order to test under relatively stable conditions some of the above conclusions, a run of the positive elements was made at Iona, September 8–9. During the time of observation and for two days previous, the wind was 0–1. A moderate haze was present at all time. The results given in Fig. 2 seem to verify these conclusions. The value of λ_{1g}/λ_1i ncreases as the conductivities decrease and it has a relatively high value at the time of maximum conductivity. Variations in n_1 lag behind like variations in the average mobility.

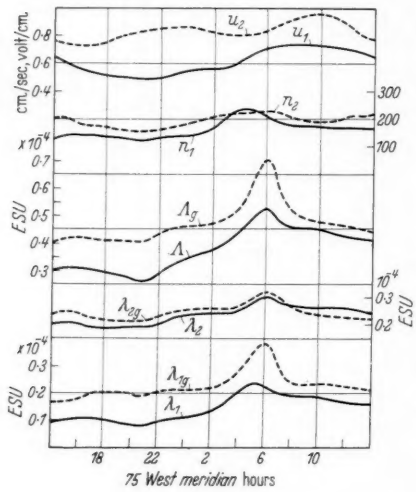


Fig. 3. Atmospheric-Electric results from Fordham.

2. *College of Mt. St. Vincent*

Conductivity measurements and observations of ion concentration were made at the College of Mt. St. Vincent from July 24 to August 8, 1950. The college is situated at the northern boundary of New York City on the east bank of the Hudson river and 15 miles from the city centre. The grounds are extensive and well removed from any densely populated area. The point of observations was 150 ft above the river level. As only day measurements were made at this location, no graphs are given. Results are summarized in Tables 1 and 2.

3. *Fordham*

Fordham University is located in a densely populated and semi-commercial district in a northern section of New York City. It has extensive grounds and is adjacent to the New York Botanical Gardens. Observations were made here in October. Fig. 3 shows the results obtained. The graphs start at 14 h, the beginning of the 24-hr periods. Due to city-air the conductivities and ion concentrations were the lowest recorded. However, the ratios of the ground conductivities to the respective 1 m values are the highest, reaching values greater than two (Fig. 11). Both the morning maximum and the night minimum conductivity occur an hour earlier than at Iona. As already noted at Iona, at the time of maximum the ground conductivity showed a greater relative variation than the 1-m conductivity. The ratios, λ_{1g}/λ_1 and λ_{2g}/λ_2 also increase for the lower values of conductivity. The positive conductivity and ion concentration show a greater variation than the negative and a much sharper rise at the time of maximum. The negative ground conductivity was generally greater than that at 1 m—an unexpected result. The calculated mobilities were found to be extremely low (Fig. 3). The number of observations at this location were not large enough to make the positive elements comparable with the negative.

VOL. 2 1952

4. *West Park*

West Park is on the west bank of the Hudson river 80 miles directly north of New York City. The district is sparsely populated. The land rises abruptly from the river in a series of hills and is generally undeveloped. The point of observation was approximately $\frac{1}{2}$ mile west of the river and at an elevation of 200 ft above it. Observations were made from November 8 to December 9. During this period a number of days were lost because of storms. North and south winds brought smoke from industrial regions. East winds gave fair to good results and west winds high values of conductivity and ion concentration. Meteorological conditions were very variable. The wind was never constant in direction and intensity for any great length of time.

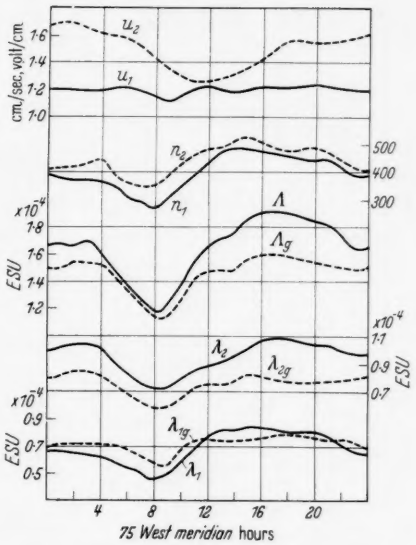


Fig. 4. Atmospheric-Electric results from West Park.

Fig. 4 summarizes the results obtained at this location. Because of the wide spread in values the vertical scale of the graphs of conductivity has been reduced to one-half of that of most other graphs. On account of the variable weather it is doubtful that any significance can be attached to the fact that the negative elements are greater than the positive. Morning fog on the river, which is common at this time of year,

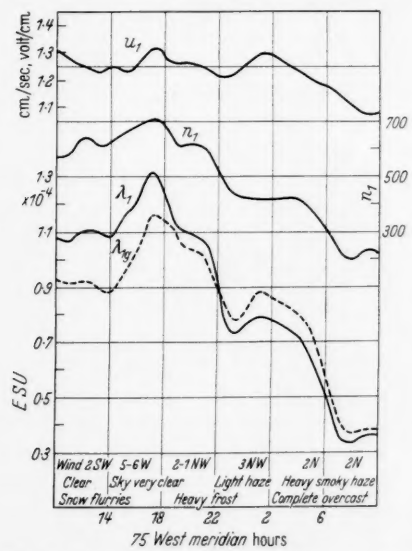

cm/sec, volt/cm
u_1
n_1
λ_1
λ_{1g}
ESU
Wind 2 SW | S-6 W | 2-1 NW | 3 NW | 2 N | 2 N
Clear | Sky very clear | Light haze | Heavy smoky haze
Snow flurries | Heavy frost | Complete overcast
75 West meridian hours

Fig. 5. Variation of positive elements with changing air conditions (Nov. 22–23).

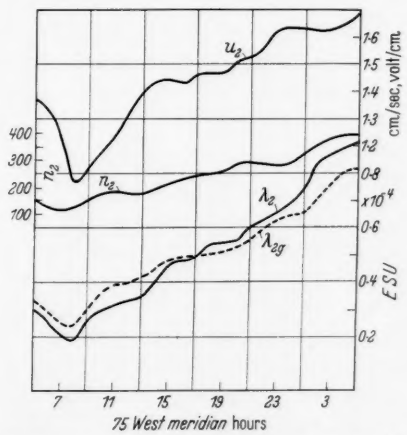


Fig. 6. Variation of negative elements with changing air conditions (Nov. 24–25).

although it did not extend to the point of observation, possibly accounts for the sudden drop in ion concentration and conductivity starting about 04 h. This phenomenon would nullify any well defined morning maximum but a secondary

minimum occurs for each of the elements a little before midnight. All ratios show a definite increase during the time of low conductivity and have their largest values between 04 h and 08 h (Fig. 11). The total 1-m conducticity was greater than the total ground conductivity which was the reverse of the results obtained at Fordham. During periods of high conductivity the average mobilities of positive ions reached values of approximately 1·50 and negative ions of 1·90 cm/sec, v/cm.

Figs. 5 and 6 are the results of two 24-hr periods of positive and negative conductivity with simultaneous measurements of ion concentration respectively. Both show progressive changes in the various elements due to changing air conditions and wind direction. Due to the particular wind direction the positive run was more variable than the negative but the negative mobilities show a greater variation than the positive.

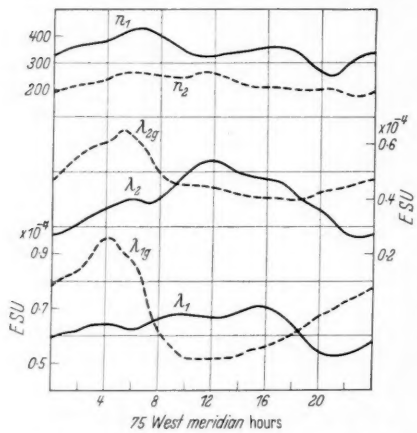


Fig. 7. Atmospheric-Electric results from Easthampton.

5. *Easthampton*

Easthampton, a small resort town, is 120 miles from New York City near the eastern tip of Long Island. This island extends from the mainland in a direction a little north of east. The mainland to the north is about 25 miles across Long Island Sound. At the point of observation the island is about 5 miles wide and very flat. Observations were made from August 15 to September 1 on a private estate 1500 ft from the ocean to the south. Conductivities and ion concentration were low during west winds which came from the main body of the island. Winds from southwest to southeast off the ocean gave higher values. No north winds were experienced. The daytime was generally sunny with light to fresh winds. At night winds were seldom greater than 1 (Beaufort scale), humidity was high and there was generally a heavy dew. A light ground mist occasionally formed in the early morning hours.

Fig. 7 and Table 3 give the results of observations at this location. Some of these results were unexpected; (1) While positive ground values of conductivity were greater than the 1-m values during the night the reverse was the case in the daytime. The change in relationship of these conductivities was always sudden but did not take place at the same time each day. λ_{1g}/λ_1 became greater than one in the evening at times varying from 18 h to 22 h and in the morning it became less than one at times between 07 h and 09 h. It was noted that the relative increase in ground conductivity coincided with the usual evening drop in wind velocity. Its decrease in the morning came at a time of increasing wind velocity. (2) The negative conductivity showed the same phenomenon but the morning change was usually later. (3) The negative ground conductivity was relatively high at night. However, the increase in the total conductivity at night, at ground level, was due more to a greater increase in positive conductivity.

All ground values of conductivity showed a definite morning maximum at approximately 04 h for the positive and 05 h for the negative. There was no sharply

defined minimum for the ground conductivities which remained relatively low and constant during the daytime and increased gradually through the night. The 1-m conductivities did show a minimum, about 22 h, and a very slight rise at the same time as the ground conductivity but the highest values were found throughout the day.

The concentration of positive small ions was always much greater than that of the negative (Fig. 7). For positive ions maximum concentration occurred between 06 h and 08 h with a minimum about 21 h. Negative ion concentration varied little

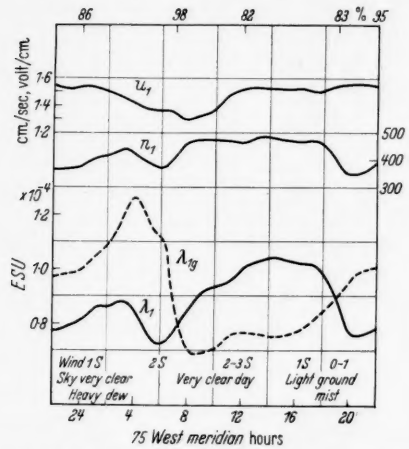


Fig. 8. Effects of turbulence and convection on positive conductivity. Relative humidity is given along the top line at the hours indicated (Aug. 17–18).

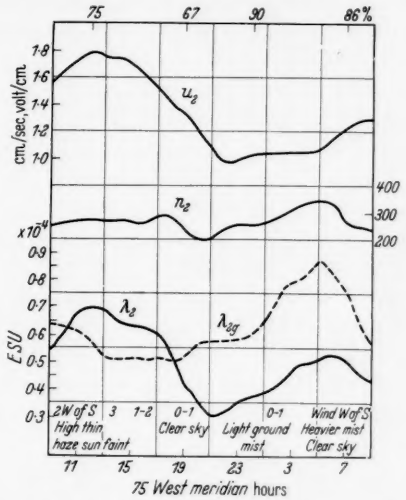


Fig. 9. Effects of turbulence convection and humidity on negative conductivity. Relative humidity is given along the top line at the hours indicated (Aug. 24–25).

Table 1. Summary of conductivities from all locations in 10^{-4} ESU

	λ_1	λ_{1g}	λ_{1g}/λ_1	λ_2	λ_{2g}	λ_{2g}/λ_2	Λ	Λ_g	Λ_g/Λ
7 a.m. to 7 p.m.									
Fordham	0·152	0·219	1·44	0·235	0·241	1·02	0·387	0·460	1·19
Iona	0·423	0·481	1·14	0·382	0·336	0·88	0·805	0·817	1·01
St. Vincent	0·443	0·453	1·02	0·456	0·410	0·90	0·899	0·863	0·96
Easthampton	0·673	0·569	0·85	0·468	0·443	0·94	1·141	1·012	0·88
West Park	0·710	0·712	1·00	0·912	0·721	0·79	1·622	1·433	0·88
7 p.m. to 7 a.m.									
Fordham	0·136	0·237	1·73	0·232	0·234	1·00	0·368	0·471	1·28
Iona	0·342	0·438	1·28	0·326	0·319	0·98	0·668	0·757	1·13
Easthampton	0·587	0·798	1·36	0·324	0·513	1·58	0·911	1·311	1·44
West Park	0·683	0·723	1·06	0·997	0·789	0·79	1·668	1·512	0·90
Mean values for each location									
Fordham	0·144	0·228	1·58	0·233	0·237	1·02	0·377	0·466	1·23
Iona	0·382	0·459	1·20	0·354	0·327	0·92	0·736	0·786	1·07
Easthampton	0·630	0·683	1·08	0·386	0·478	1·20	1·026	1·156	1·13
West Park	0·696	0·717	1·03	0·954	0·755	0·79	1·650	1·472	0·89

remaining fairly constant between 05 h and 12 h and falling gradually to a minimum value at 22 h. Average mobilities of both positive and negative ions were more variable than at any other location. Variation between high values in the day and low values at night was considerable (Table 3).

Table 1 is a summary of all conductivity elements from the various locations. The results are shown in three groups; the mean for each location for the day, night and all observations. The locations are arranged in order of increasing conductivity at 1 m. The ratios given decrease in going from the location of lowest to that of highest conductivity, except some of those from Easthampton. During the day the total 1-m conductivity was greater than the total ground conductivity, except at Fordham (location of lowest conductivity) but at night the total ground conductivity became the greater, except at West Park (location of highest conductivity). Values of the ratios for the night are greater than their corresponding day values. Easthampton showed the greatest increase in night ratios over day.

Table 2 is of ion concentrations and various ratios with the locations arranged in the same order as in the first table. With few exceptions this is the order of increasing ion concentrations of both signs. In general the mobilities increase from the location of the lowest ion concentration to that of the highest. However, the mobility of negative ions for Easthampton, shows a very large decrease at night.

Table 2. Summary of ion concentrations and mean mobilities

	n_1	n_2	n_1/n_2	λ_1/λ_2	u_1	u_2
			7 a.m. to 7 p.m.			
Fordham	161	196	0·82	0·65	0·66	0·84
Iona	260	195	1·33	1·11	1·14	1·37
St. Vincent	258	231	1·12	0·97	1·22	1·38
Easthampton	356	235	1·51	1·45	1·34	1·40
West Park	414	462	0·89	0·78	1·20	1·37
			7 p.m. to 7 a.m.			
Fordham	169	190	0·89	0·59	0·56	0·85
Iona	214	163	1·31	1·05	1·12	1·40
Easthampton	341	217	1·57	1·81	1·22	1·05
West Park	396	434	0·91	0·70	1·20	1·59
			Mean values for each location			
Fordham	165	193	0·85	0·62	0·61	0·84
Iona	237	179	1·32	1·08	1·13	1·39
Easthampton	349	226	1·54	1·63	1·28	1·23
West Park	405	448	0·91	0·73	1·20	1·49

Discussion of Observations

Conductivities at ground level and at 1 m

Because of their "abnormal character," the Easthampton results are not included in the following discussion. They will be discussed separately.

As the ratios of the polar and total conductivities at ground level to the respective 1-m conductivities showed a definite variation at individual locations and a greater one among the locations, causes must be local conditions and general meteorological

conditions. Results have shown that the 1-m polar conductivities are more variable than ground values, except at the time of maximum conductivity, and this is of course reflected in the total conductivities. Thus, the above ratios usually varied in the opposite sense from conductivity. This phenomenon should be expected when we consider that the effects of convection and turbulence are greater at higher levels than at ground. The mean eddy coefficient as given by W. SCHMIDT [8] increases between 11 cm and 244 cm from 0·26 to 3·40 cm^{-1}, g, sec^{-1}. Fig. 10 which is made up of 371 pairs of positive conductivity at 1 m and at ground level, shows the effect of increasing turbulence on the ratio, λ_{1g}/λ_1. Increasing values of λ_1 were so chosen that they corresponded approximately to increasing wind velocity and averaged in groups between the values indicated in the figure. The ratio, λ_{1g}/λ_1 obtained from the average values of λ_1 and the averages of λ_{1g} was plotted against λ_1. The simultaneous pairs of values were taken from all locations except Easthampton and there is of course not an equal distribution of values among the points. The broken line gives λ_{1g} plotted against λ_1. The average increase in λ_1 and λ_{1g} from point to point along the three superimposed straight lines on the curve expressed in percentages, is as follows:

	Line 1	*Line 2*	*Line 3*
λ_1	67 %	22 %	10 %
λ_{1g}	37 %	15 %	9 %

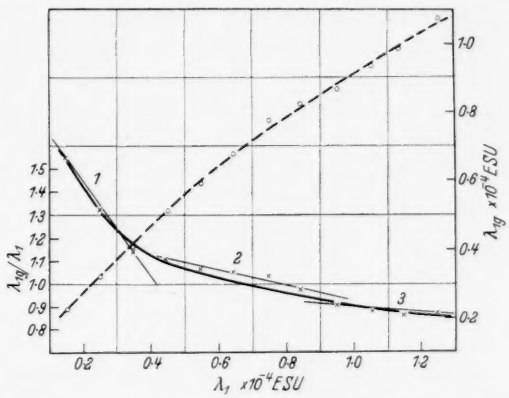


Fig. 10. Variation of λ_{1g}/λ_1 with increasing λ_1 and wind velocity.

It can be seen that λ_1 increases faster than λ_{1g} although at a decreasing rate. The initial sharp drop in the curve would indicate that no great turbulence is necessary to influence the 1-m level but there is a considerable lag in its effect at ground level.

Local factors that cause variation in ion concentration and indirectly in conductivity such as a change in the amount of radioactive matter in the air, a variation of condensation-nuclei or in air pollution, or any combination of these are also airborne. Therefore, the metre level should be affected before the ground.

An additional mechanism that could account in part for the fact that the 1-m conductivities are more variable than those at ground is that q, the rate of formation of small ions, not only has different values at ground and at 1 m but it may also vary at the two levels in a different manner. The average rate of production of small ions near the surface of the earth over land is due almost entirely to the radioactive matter in the air and in the earth, cosmic rays contributing only a small fraction of the whole. Approximately 30 % of q up to 1 m above ground is due to gamma radiation from the earth itself as shown by HESS [9], [10]. Very close to the surface beta radiation from the earth would play an increasing part in ionization. Experiments have been carried out by Dr. HESS and the author with two identical ionization chambers, one at the metre level and the other a few centimetres from the

VOL. 2 1952

ground. Results indicate an appreciably increased rate of ionization at ground level over that at 1 m. Other experimenters have obtained similar results ([3], [7]). For a given location, the rate of ion formation due to the radioactive matter in the earth can be considered constant as long as the condition of the soil remains unchanged. This should be true even though the amount of radioactive matter, of pollution or of condensation-nuclei in the air near the surface was changing. Thus, the rate of ion formation due to the radioactive matter in the earth only, might be considered a parameter having a larger value at ground than at 1 m. Any change in total q at ground and at 1 m, even though numerical equal at both levels, would be relatively less at ground as the rate of ion formation at ground level contains a larger constant part.

In the normal electric field of the atmosphere, λ_{1g} should be greater than λ_1 due to a greater concentration of positive ions near the ground and λ_{2g} should be less than λ_2 due to a depletion of negative ions near the ground. The mean values of λ_{1g}/λ_1 are greater than one (Table 1) except for the day value at West Park. As the smallest value of λ_{2g}/λ_2 was also found at West Park in the day, the cause of this exception is local. The local effect could be the morning fog over the river and the fact that smoke when present was more prevalent in the daytime. The sources of both these factors were below the point of observation and apparently affected ground level before the metre level to a degree that tended to predominate over other factors which could cause an increase in ground conductivity over the 1-m value. Results show that λ_{2g}/λ_2 was greater than one only at Fordham. The ratio, λ_{1g}/λ_1 was also greatest at this location. Johnson and Wait [7] from the results of their Atmospheric-Electric Survey in Washington conclude that atmospheric pollution may be a large factor in producing low conductivities. As our largest values of conductivities were found at West Park when the air was clear and our lowest in city-air at Fordham, our results add weight to this view. This factor would usually be more effective (discussed above) in reducing conductivity at 1 m than at ground level. This with the fact that the rate of ion formation is greater at ground level could cause the ratio, λ_{2g}/λ_2 to become greater than one where conductivity is low.

If the electric field at ground does not differ appreciably from that at 1 m, as a number of experimenters have demonstrated, then it is generally considered tha Λ should equal Λ_g. Our results show that they are not equal, although the values of Λ_g/Λ do not differ widely from unity except where conductivities are very low. It would be entirely fortuitous that factors which cause unequal variation in the polar conductivities should act in such a way as to make the sum of the polar conductivities at ground level equal to the sum of these elements at the metre level. Therefore, for the same reasons that the polar conductivities vary, total conductivities should show some unequal variation. Fig. 11 shows that the variation of Λ_g/Λ is similar to that for the polar conductivities.

The values of Λ_g/Λ for West Park are probably not entirely significant except to show diurnal variation. Although a large number of observations were made at this location, meteorological conditions were so variable from day to day that Λ and Λ_g are hardly comparable. At Fordham the number of observations were too few to make the polar conductivities at the two levels comparable. However this adverse factor should practically cancel out in the ratio, Λ_g/Λ, as all the values

of λ_{2g}/λ_2 at this location were very close to unity and meteorological conditions varied little for the period of observation.

As is evident from the various figures, the ground conductivity shows a relatively sharper increase approaching and at the time of maximum than the 1-m conductivity. McNish and Wait [11] report that the diurnal maximum of q occurs approximately at sunrise. A minimum of condensation-nuclei is also found in the early morning ([7], [12]). Yunker [13] has found a considerably lower minimum at 50 cm than at 20 m above ground. From this we might conclude, that before convection has become pronounced, the rate of ion formation is a maximum and greater at ground level than at 1 m and this combined with the fact that the loss of small ions due to condensation-nuclei would be less at ground, should account for the sharper rise in maximum ground conductivity.

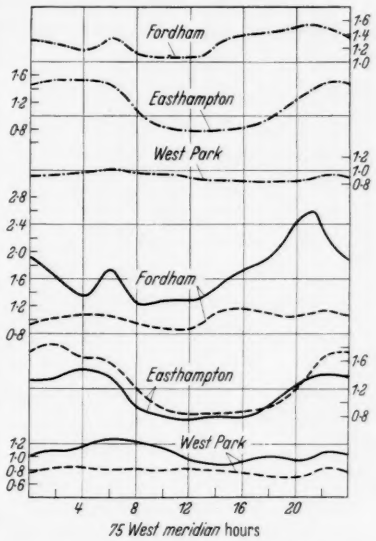


Fig. 11. Diurnal variation of $\frac{\lambda_{1g}}{\lambda_1}$, $\frac{\lambda_{2g}}{\lambda^2}$ and $\frac{\Lambda_g}{\Lambda}$.

Discussion of the Easthampton results

Table 3 summarizes the results obtained at Easthampton. The third line gives the percentage change in values of the ion concentrations and conductivities at night over those in the day. The table shows that the positive elements at both levels, day and night, predominate over the negative. Humidity which was greater at night influenced the negative conductivity adversely as will be shown in the discussion of ion concentration and average mobility.

Fig. 8 and 9 give the results of one 24-hr run of positive conductivity and one of negative respectively. An increase in the ground conductivities and a decrease in the 1-m values are associated with a drop in wind velocity as indicated in the figures. The early evening hour at which this occurred would also be a time of decreasing convection. In the morning a reversal in the relationship of the ground and 1-m conductivities follows a rise in wind velocity. In the case of the negative run the reversal occured at a later morning hour (10 h) due to

Table 3. Atmospheric-Electric results from Easthampton

	λ_1	λ_{1g}	λ_2	λ_{2g}	Λ	Λ_g	n_1	n_2
Day	0·673	0·569	0·468	0·443	1·141	1·012	356	235
Night	0·587	0·798	0·324	0·513	0·911	1·312	341	217
Percent change	−13	40	−31	15	−20	30	−4	−8

	λ_1/λ_2	$\lambda_{1g}/\lambda_{2g}$	n_1/n_2	λ_{1g}/λ_1	λ_{2g}/λ_2	Λ_g/Λ	u_1	u_2
Day	1·44	1·28	1·51	0·85	0·94	0·88	1·34	1·40
Night	1·81	1·55	1·57	1·36	1·58	1·44	1·22	1·05

VOL. 2 1952

delayed convection as the sun was obscured by a high haze until about 11 h. Here maxima and minima are more pronounced than in the mean of all results.

Advantage was taken of a night of brisk winds, August 28–29, to check the influence of wind velocity on λ_{1g}/λ_1. Observations were made from 16 h to 07 h the following morning. For the whole time wind was 4–6 S from the ocean which was very rough. The winds were accompanied by low thin clouds. Results of 20 pairs of 1-m and ground conductivities were $\lambda_1 = 1{\cdot}057$, $\lambda_{1g} = 0{\cdot}862$, $\frac{\lambda_{1g}}{\lambda_1} = 0{\cdot}82$ and the mean of the average mobilities was 1·28 cm/sec, v/cm. The usual reversal to higher values of ground conductivity did not occur. Turbulence, at least at this location, is an important factor in determing the relation of the polar ground conductivities to the respective 1-m values.

Scrase [14], using a Wilson apparatus, found large negative charges entering the ground against a positive field. P. J. Nolan [4] also measured this "non-ionic negative current." This phenomenon was associated with fog or mist, a high potential gradient, and high humidity (Kew—98–100 %; Glencree—80 %). The relative humidity at midnight over the period of observation at Easthampton had a mean value of 89 %. Scrase attributed this negative current to the settling down under gravity of very small droplets having negative charges. If this phenomenon occured at Easthampton, our apparatus would not measure the charge of these droplets because of their low mobility but would record the negative ionic conductivity. We seemed to have observed at this location, as the wind was generally from the ocean, an extreme example of the Lenard Effect which is characterized near surf by a high potential gradient, positive space charge and droplets of water carrying negative charges. The variation in conductivity between day and night is then simply due to greater convection and turbulence and lower relative humidity (midday mean—63 %) in the day superimposed upon factors producing the Lenard Effect and that cause diurnal variations.

Polar conductivities and ion concentrations at 1 m above ground

The diurnal variation in the concentration of small ions at the different locations, as expected, has the same character as that for the conductivities although the variation are not always in phase. Tables 1 and 2 show that the polar conductivities increase with the ion concentrations of the same sign. However, the ratios of these conductivities to the simultaneous ion concentrations, and thus the average mobilities, are not constant. Usually average mobilities were high when both conductivities and ion concentrations were high and low when these elements were low (Table 2). This would indicate that the factors that govern the concentration of small ions other than those which determine their rate of formation must have some effect on their mobilities.

At West Park the mean value of λ_1 decreased at night while λ_2 increased but the ratio, n_1/n_2 remained nearly constant day and night. The increase in λ_2 during the night appears due to an increase in the mean mobility of the negative ions for the night period (Table 2).

At Easthampton the mean value of λ_2 was 31 % less at night than in the day but n_2 was only 8 % less (Table 3). Therefore, the decrease in negative

conductivity at night is due to a considerable decrease in the mean mobility of the negative ions.

A better insight into the relationship of conductivity to ion concentration can be obtained by a study of a few 24-hr periods of observations. Fig. 5 which shows a progressive decrease in positive conductivity and ion concentration shows a similar decrease in the mobility of the ions. The mean mobility for the 24 hrs was 1·22 cm/sec, v/cm. For values of conductivity (at 1 m) greater than $1{\cdot}00 \times 10^{-4}$ ESU, the mean was 1·26 when the air was clear. For values of λ_1 less than 1·00 when a light to heavy haze was present the mean was 1·19 and for values less than 0·400 when a heavy smoky haze had formed the mean became 1·09.

Fig. 6 shows a continual increase in the negative elements as the air became clearer. The mean mobility for the 24 hrs was 1·44. For values of λ_2 less than 0·400 was the air was very smoky the mean mobility was 1·28, for values between 0·400 and 0·600 observed when the air was clearer due to a change in wind direction the mean mobility was 1·44 and for values of conductivity greater than 0·600 after the haze had entirely disappeared and the air was very clear, the mean became 1·62. Humidity varied little during the above runs. This would indicate that changes in the smoke content of the air cause parallel changes not only in the concentrations of both positive and negative ions but also in their average mobilities.

VOL. 2 1952

Fig. 8 and 9 give two separate days of observation made at Easthampton on the dates indicated. For both runs the wind was from the ocean. The negative mobilities in Fig. 9 show a considerable variation. The lowest values occurred during the night hours when the relative humidity was high and a light ground mist was present. The day and night mean values were 1·56 and 1·07 respectively. During this particular run the air was clear, and except at the beginning and end of the period the sky was clear. Smoke was never prevalent at any time at this location. The positive elements in Fig. 8 were observed under particularly good atmospheric conditions with very clear air for the whole period of observation. Approximately the same cycle of changes occurred as for the above negative elements. Mobilities were constant except in the late morning hours. These lower values did occur at the time of highest relative humidity but the positive mobility was always found to be lowest at this time and seems to be due to the fact that this is the time of "transition" when the ground and 1-m conductivities reverse their relative values. Approaching the second midnight of this run humidity again became high (95%) but mobilities showed no decrease. A run of the positive elements made August 22–23 had about the same relative variation in mobility of the ions as those in Fig. 8, but the values were lower. The mean for the daytime was 1·33 and for the night 1·25. The lower values were due to the fact that the wind direction was not as favorable being partially off shore. However, the ralative humidity was low (lowest for the month of August). The value at midday was 64% and at midnight 76%. These observations indicate that relative humidity is not a contributing factor in causing variations in the mobilities of positive ions and indirectly in positive conductivity but it does influence negative conductivity.

The mobilities of both positive and negative ions determined from the Fordham results (Table 2) were extremely low. Atmospheric conditions at Fordham and Iona were not very different for the respective periods of observation, yet the mobilities

for the latter location were a good deal higher. The mean relative humidity at noon and at midnight for the two months of observation at Iona were 59% and 74% respectively and for Fordham the corresponding means were 61% and 76%. This should eliminate relative humidity as one of the causes of the difference in mobilities at the two locations. Johnston and Wait [7] found a much higher concentration of condensation-nuclei in Washington, D.C., and near suburbs than at more distant locations within 20 miles of the city. The lowest value of conductivity was found by them at a place where the concentration of nuclei was greatest. A very high concentration of nuclei should be expected in New York City, especially if we accept the view that such nuclei are primary products of combustion. Sources would be not only industrial and domestic fires but the tremendous volume of motor traffic in the city. Yunker [15] has shown that the number of ions per mobility group varies inversely as the density of atmospheric condensation-nuclei. His results indicate that not only does an increase in the number of condensation-nuclei cause a transfer of ions to lower mobility group but also causes a lowering of the mean mobility in a given range. Therefore, at Fordham the low mobilities of the positive and negative ions with respect to other locations must be due to a high concentration of condensation-nuclei.

The results show that the polar conductivities are dependent on the concentration of the respective ions and their *average mobilities* both of which seem to be affected by the same changing conditions. There is some evidence that, under certain local conditions, variations in mobility is the more important factor in determining the diurnal variation in conductivity (Fig. 9). The diurnal variations of conductivity and ion concentration are seldom in phase because of variations in the average mobilities of the ions.

We have seen that the polar conductivities at the ground and metre levels are not influenced to the same extent by atmospheric and local conditions. This would mean that the concentration of small ions and their average mobilities, which determine conductivity, are not equally affected. Thus it can be assumed that variations in ion concentration and mobility at the two levels would differ in degree and possibly in phase. An appreciable difference might form between the number of small ions of opposite sign at a given level or between small ions of the same sign at different levels. This need not mean that the total number of ions, large and small, is changed a great deal. Considerable evidence shows that the electric field is about the same at ground level and at 1 m. A study of the equation,

$$p = [(N_1 - n_2) + (n_1 - N_2)]$$

which gives the relationship between the space charge and the concentrations of large and small ions, shows that any recombination of ions of opposite sign or the agglomeration of small ions on uncharged nuclei would leave the space charge unchanged. Thus the number of small ions might vary considerably without a notable change in the difference between all positive and all negative ions. Conductivity of one sign is a function of the concentration of small ions of the same sign and their average mobilities while the electric field depends on space charge in which large ions usually predominate over small ions. Therefore, while the former might vary unequally at two levels the latter could remain constant.

Acknowledgments—The author wishes to thank Dr. T. G. BULLEN of Iona College who assisted in making observations at all the locations and FR. BURNS, S. J. and Mr. PAUL KENNY who helped with observations at Fordham.

Special thanks are due to Dr. V. F. HESS for his valuable guidance and helpful advice during the progress of the work.

REFERENCES

[1] WATSON, R. E.; London Meteor. Off., Geo. Mem., No. 45 (1929. [2] SCRASE, F. J.; London Meteor. Off., Geo. Mem., No. 60 (1934). [3] HOGG, A. R.; Com. Solar Obser. Mem., No. 7 (1939). [4] NOLAN, P. J.; Proc. Roy. Irish Acad. 1940 **46** A **6** 65–75. [5] WAIT, G. R.; Terr. Mag. 1927 **32** 31–36. [6] JOHNSTON, H. F. and WAIT, G. R.; Terr. Mag. 1931 **36** 32–40. [7] WAIT, G. R.; Terr. Mag. 1931 **36** 111–131. [8] SCHMIDT, W.; Probleme der Kosmischen Physik, Hamburg, Henri Grand, 1925 **7** 88. [9] HESS, V. F.; The Electrical Conductivity of the Atmosphere. Constable and Co., London 1928. [10] HESS, V. F.; Archiv für Meteor., Geophysik 1950 **3** 56–63. [11] MCNISH, A. G. and WAIT, G. R.; Phys. Rev. 1935 **47** 789. [12] BRADBURY, N. E. and MEURON, H. J.; Terr. Mag. 1938 **43** 231–240. [13] YUNKER, E. A.; Terr. Mag. 1940 **45** 121–126. [14] SCRASE, F. J.; London Meteor. Off., Geo. Mem., No. 58 (1933). [15] YUNKER, E. A.; Terr. Mag. 1940 **45** 127–132.

VOL.
2
1952

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 216 to 225. Pergamon Press Ltd., London

An explanation of radio propagation at 16 kc/sec in terms of two layers below *E* layer

R. N. Bracewell* and W. C. Bain**

(Received 8 December 1951)

ABSTRACT

An ionospheric model consisting of two ionized layers below the normal *E* layer is postulated in an attempt to account for observed features of the propagation of 16 kc/sec radio waves over distances of 90 and 535 km. These observations suggest the existence of more than one ionized layer below the *E* layer.

The lower of the two postulated layers reflects very long waves over the 535 km range, but over 90 km this layer is penetrated and the waves are reflected from the upper layer. The field produced by a transmitter at distances up to about 300 km is composed of the ground wave plus waves reflected one and more times from the upper layer. At 535 km, waves from both layers contribute to the total field, especially the wave reflected once from the lower layer and waves reflected more than once from the upper layer.

New experimental data for 535 km are interpreted in terms of this model.

Introduction

New information about the lowest parts of the ionosphere can be of immediate use in several branches of atmospheric physics, and in solar physics. Although very little is known about the ionosphere below *E* layer, many interesting phenomena take place there. Theoretical work concerned with the base of the ionosphere is at present handicapped by lack of facts. This is particularly the case, for example, with the vigorous investigations aimed at elucidating the processes by which the ionization is produced, and with investigation of the spectrum and places of origin in the sun of the ionizing radiation.

The echo technique so successfully used in exploring the more familiar ionospheric layers, is only slowly becoming available at the low radio frequencies which are employed in studying the base of the ionosphere. A suitable technique of a different kind has been developed in Cambridge, which has as its essential feature the measurement of phase changes in the sky waves received from C.W. transmitters. The application of this technique has led to the publication of a considerable body of new facts, on which the present paper is based.

The material utilized in this paper is contained mainly in papers by Straker [1], Bracewell [2], and by Bain, Bracewell, Straker and Westcott [3], which respectively give accounts of the propagation of 16 kc/sec waves over distances of 90, 200 and 535 km. Observational material obtained subsequently at Aberdeen is described and used. Further material is taken from a paper by Weekes [4], which describes measurements made in an aircraft flying outwards from the transmitter to distances of 800 km.

* Division of Radiophysics, Commonwealth Scientific and Industrial Research Organization, Australia.

** Department of Natural Philosophy, University of Aberdeen, Scotland.

We are concerned throughout with the 16 kc/sec transmission from GBR Rugby. The chief receiving points were Cambridge at a distance of 90 km, and Aberdeen, at a distance of 535 km. The symbols and terms used throughout are the same as in the previous papers, and are fully explained in [2].

I. Nomenclature

In this paper we shall postulate the existence of two distinct banks of ionization below the E layer. According to the established system of naming ionospheric layers, two layers below E layer should be named C and D or D_1 and D_2. We do not yet know enough about the structure of the base of the ionosphere to allot these symbols confidently. They should be reserved until there is reasonable agreement about the facts. Until then, as already the literature bears witness in respect of all four of the above symbols, their significance will vary from author to author, and confusion may result. We propose therefore, in discussing the two layers dealt with here, to adopt provisional symbols *outside* the established system. The future allocation of C, D, D_1, D_2 is thus not prejudiced. $D\alpha$ and $D\beta$ will be used here, $D\alpha$ referring to the higher layer.

II. The Two Layer Model of Very Long Wave Propagation

We know that at 90 km from GBR Rugby the ionospheric wave consists almost entirely of a single downcoming ray [4], [1]. At distances of about 200 km we find [2], in addition, multiple reflections *from the same layer*. The diurnal movement of the reflecting level begins at sunrise and then closely follows the formula

$$h = 72 + 5{\cdot}5 \log \sec \chi \text{ km},$$

where h is the "apparent height" defined by Bracewell, Budden, Ratcliffe, Straker and Weekes [5], no phase change being allowed at reflection, and χ is the Sun's zenith distance. More refined observation reveals that (a) a small fall in apparent height precedes the main drop at sunrise [1], and that (b) the variation lags the formula slightly so as to be symmetrical about a time 25 min after local noon [2]. The upper curve of Fig. 1 shows the height variation deduced from observations made at a distance of 90 km, at the equinoxes [1].

Now at 535 km one observes a steady pre-sunrise fall in height, which is considered in [3] to be about 18 km. Though a full interpretation of the observations was not attempted there, it was definitely concluded that a ray was present which was reflected once from a level which fell steadily through a great height, before sunrise. This conclusion is illustrated diagrammatically by the lower curve of Fig. 1. The fall in height is maintained from the time when the Sun's zenith distance is 98 degrees until approximately sunrise, with an average rate of 3 km per degree change in χ. In the evening, the recovery begins about sunset and is usually almost complete when χ reaches 98 degrees; any systematic lag in recovery does not exceed 2 degrees of χ, and may be much less. A lag of 2 degrees would mean symmetry about a time 6 min after noon. The height variation between sunrise and sunset is not yet known exactly but it is known to be small and is represented in Fig. 1 as constant.

These observations seem incompatible with what would be expected if reflection were from the same layer at both 90 and 535 km. Thus over a period when the height of reflection at steep incidence is practically constant, the height at more oblique

incidence falls 18 km. Furthermore, the height variation at steep incidence is assymmetrical about noon, whereas at more oblique incidence it is relatively symmetrical.

So much for height variation. As regards the reflection coefficient $_{||}R_{||}$*, it is found that reflection at 90 km during the day-time is strong in winter ($_{||}R_{||} = 0{\cdot}4$), weaker in summer (0·12), and practically constant from sunrise to sunset on a given day [1], [2]. The reflection coefficient over 535 km during twilight is the same in all seasons, approximately 0·5 [3], and must be much the same during daylight.

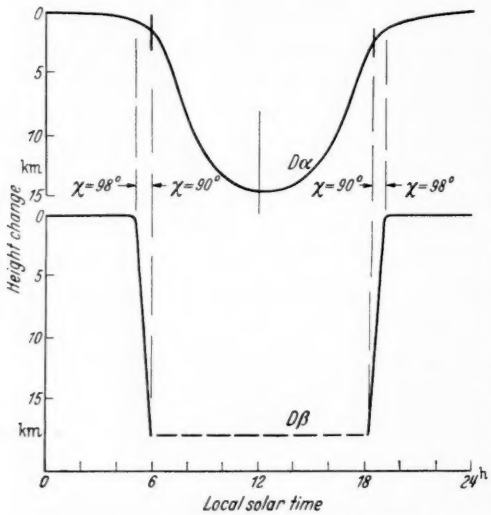


Fig. 1. The diurnal height variations of the layers $D\alpha$ and $D\beta$.

From the foregoing résumé of observational data it does not appear possible to give an explanation of the field variations at 90 and 535 km in terms of one simple ionized layer. We postulate a two-layer model as the next simplest hypothesis; but greater complexity is not excluded by the evidence.

The layer which is responsible for reflection over 90 km at 16 kc/sec we shall call $D\alpha$. We postulate the presence of a lower layer, to be called $D\beta$. Only the daylight and twilight are considered. The existence or behaviour of the layers during the night when $\chi > 100$ degrees, will not be discussed. To complete the specification of the two-layer model, we assume $D\alpha$ and $D\beta$ to vary in height as in Fig. 1. Nothing is assumed about the reflection coefficient of $D\alpha$, but $D\beta$ is taken to be penetrable at steep incidence and strongly reflecting at more oblique incidence. For definiteness it will be assumed that $D\beta$ has constant properties from sunrise to sunset.

On this model there must be a critical range or zone where one layer takes over from the other. Observations of GBZ Criggion (15·2 kc/sec) made at Cambridge, and of GBR Rugby at Edinburgh show morning phase onsets respectively similar in timing to GBR observed at Cambridge and at Aberdeen. Hence, in these latitudes, the changeover occurs between 222 and 419 km from the transmitter.

The two layer hypothesis has been formulated to account only for the limited data summarized above. We shall next examine the hypothesis in the light of published observations and then show how it can be used to explain certain detailed but hitherto unpublished observations of the Rugby signals received at Aberdeen.

III. Implications of the Two Layer Model Compared with Observation

The most obvious implication of the hypothesis is that the absolute heights of reflection from the two layers would probably be different. Unfortunately absolute reflection heights at these long wavelengths have not been determined with precision.

* The subscript signs in the standard symbol $_{||}R_{||}$ indicate that it is a ratio of two electric fields both of which are parallel to the plane of propagation.

The most that can be said is that the noon height is about 70 km, with an uncertainty of about 8 km, at both short and long distances. Hence from this evidence alone, the heights of reflection at noon might differ by up to 16 km.

It is also to be expected, from the hypothesis, that the ground interference pattern observed on going away from the transmitter should show a discontinuity at the distance at which $D\beta$ takes over from $D\alpha$, *i.e.* between 222 and 419 km. Confirmation of this expectation is found in the recent work of WEEKES [6], who finds a "change in reflection conditions between 300 and 500 km."

In its present form the hypothesis has not involved the electron densities in the two regions. Ionization densities are not given by long wave measurements of the type concerned. Similarly we have not included a postulate as to polarization of the reflected wave at the greater distances because the available observations are inadequate to establish it. Indeed the necessary observations will probably be difficult because, as we shall see later, there is evidence of multiple reflections and in this case interference between the rays will lead to measurements of the polarization of total field which is not simply related to the polarization of individual rays. Thus to take a particular example, the observations of 3 May 1949, given in [3], showed that the phase difference between the normally and abnormally polarized components of the total field varied smoothly from 0 up to 240 degrees between 5·30 a.m. and noon. Similarly WEEKES' ground interference pattern shows a transition from elliptical to linear polarization of the total field between 335 and 535 km. Both these effects appear due to interference between multiple reflections.

VOL. 2 1952

Aberdeen loci: Loop in plane of propagation

In this and the following sub-section a detailed comparison is made between behaviour predicted by our two-layer model and certain detailed observations of the field of GBR Rugby observed 535 km away at Aberdeen.

According to our hypothesis, the field at Aberdeen can be described as follows.

The once reflected wave comes from $D\beta$. The twice reflected wave makes a first hop of 268 km, which may be deemed to be less than the critical range, and propagation is therefore via $D\alpha$. Reflections of higher orders are from $D\alpha$. This scheme is illustrated in Fig. 2. To test the scheme, we can compare its implications with new observational material and with details reported in [3].

In Fig. 3 we see experimental loop induction loci on the complex plane, representing the signal received by a loop aerial in the plane of propagation in phase and amplitude. In Fig. 3, (a) refers to a winter and (b) to a summer morning. Both loci are approximately the same until a little before the Sun rises. From sunrise on it is exprected that the field variations will be caused mainly by the changing phase of the wave reflected twice from $D\alpha$.

This ray would trace out a circular locus whose radius $2H_1^{(2)}$ relative to the ground wave H_0 would be given by*

$$\frac{2H_1^{(2)}}{H_0} = 3{\cdot}9\,|R_{||}|^2 = \begin{matrix} 0{\cdot}62 \text{ (winter)} \\ 0{\cdot}06 \text{ (summer)} \end{matrix},$$

* H_1 is the horizontal magnetic component of the downcoming ray, but because of ground reflection the quantity observed is $2H_1$. The superscript in parentheses gives the number of hops.

where the factor 3·9 is calculated from the transmission geometry and includes ground wave attenuation. Thus in the summer the variation after sunrise should be practically negligible, which agrees with observation. The loops observed on winter mornings are in reasonable agreement with calculation. Using the formula $2H_1^{(3)}/2H_1^{(2)} = 1{\cdot}6\ {}_{||}R_{||}$ to calculate the strength of the wave reflected three times, we find it to be as high as 0·64 of the twice-reflected in winter, and therefore important. It would alter the expected circular locus into a closed epitrochoid with a single loop. The day to day variability of the observed loci is too great to allow this fine detail to be definitely recognized.

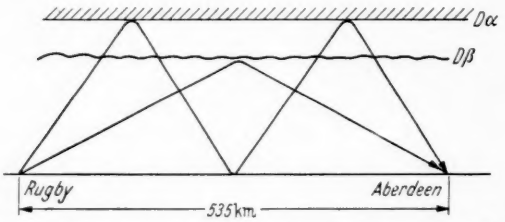


Fig. 2. The main rays composing the sky wave field at 535 km.

The reflection coefficient of $D\alpha$ is a minimum in August and increases rapidly in September and October, thus departing from strict symmetry about the equinoxes [1], [2]. The size of the daylight loop in different seasons varies similarly. The terms summer and winter will be used to refer respectively to the seasons in which $D\alpha$ reflection is weak and strong.

The size of the daylight part of the locus is thus considered to be compatible with prediction from the model. It is not compatible with what would be expected of a possible ray reflected once from $D\alpha$. In this case the same high reflection coefficient would be required in winter as holds at steep incidence, in spite of the already strong reflection from the lower layer. On account of the single reflection, the seasonal variation in size of the daylight part of the locus would be only 3:1 instead of 10:1, which should produce in summer a loop of readily observable size. Finally, the twice reflected ray should amount to 0·6 of the once reflected ray in winter and rotate at $3\frac{1}{2}$ times the speed, thus producing marked epitrochoidal loops at intervals of $360/(3\frac{1}{2}-1) = 140$ degrees, which could hardly fail to be observed.

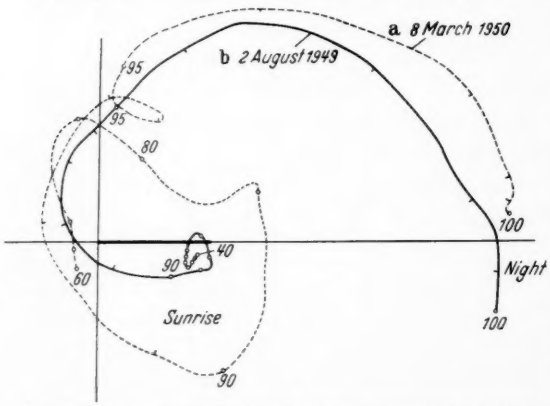


Fig. 3. Loop induction loci on the complex plane, taken with the loop in the plane of propagation, showing summer (broken line) and winter (full line) types of behaviour. Parametric values of χ are indicated.

A further test results from examining the total amount of phase change between sunrise and noon. In winter we expect a total change of about 400 degrees of the ray reflected twice from $D\alpha$. On four occasions studied, the amount seemed rather less than one cycle. However, it is not clear what centre should be chosen for this purpose, and it is known that large disturbances are present due to waves reflected three and more times from $D\alpha$. Agreement with the model is thus not unreasonable, but not close. It is found that matters are improved if it is supposed that the small pre-sunrise fall in height observed at 90 km, is more marked at more oblique incidence.

Thus far only the morning has been considered. In the afternoons a curious asymmetry was found in some seasons. Fig. 4 shows typical diurnal curves of signal strength, taken with a loop aerial in the plane of propagation, in summer and winter. In summer, accurately symmetrical curves were usual, but in winter very strong asymmetry was usually observed. The deep sunrise dip in strength would often be completely absent in the afternoon. No seasonally dependent asymmetry was noticed in the experiments at shorter ranges. The only diurnal asymmetry was the small afternoon lag in the ascent of $D\alpha$, present in all seasons. This phenomenon appears however, to be sufficient to explain the Aberdeen observations.

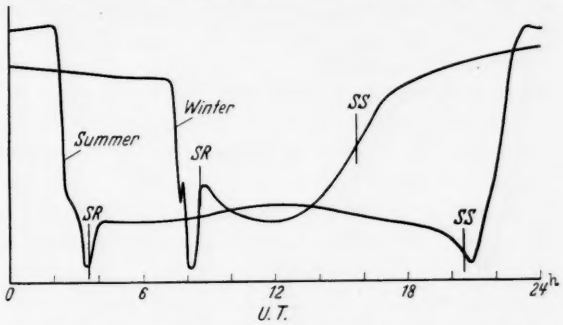


Fig. 4. Signal strength records made with a loop aerial in the plane of propagation, showing the asymmetry often observed in winter but not in summer. (SR = sunrise, SS = sunset.)

First we note that in summer the contribution via $D\alpha$ is negligible during daylight hours, being only 0·06 of the ground wave. We therefore expect to observe the high degree of symmetry associated with the variations of $D\beta$ until after sunset when reflection from $D\alpha$ begins to strengthen.

In winter it is expected that the morning locus will be retraced in the afternoon, but with the usual lag associated with $D\alpha$. By the time the Sun sets, recovery to the sunrise point is thus not completed. Nevertheless the recovery of $D\beta$ then sets in, preventing the remainder of the morning loop from being traced out at all. Fig. 5a (full line) shows a winter-type locus observed on the morning of 5 April 1950. There was the usual deep sunrise minimum A, followed by a typical daylight loop. The broken line is a constructed afternoon locus, calculated from the morning observations by splitting into two parts at sunrise, and recombining after allowing the proper time lag in the daylight part. It is seen that by sunset the morning loop is only partly retraced. The rest of the recovery of $D\alpha$ is superimposed

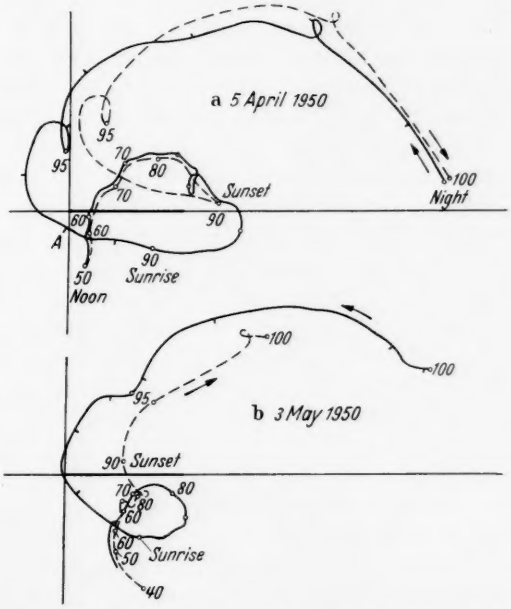


Fig. 5. Winter-type loop induction loci exhibiting the nature of the diurnal asymmetry on the complex plane: a) observed morning and calculated afternoon, b) observed morning and afternoon. Parametric values of χ are indicated.

on the recovery of $D\beta$. There is no deep minimum corresponding to that in the morning. The afternoon was not observed on the day in question, but observations on other days confirm the explanation inasmuch as sunset is found to fall on the side of the daylight loop opposite to sunrise. After sunset a recovery to night conditions is found, often along a locus displaced from that traced during the morning twilight. Such a displacement is required by the explanation. Fig. 5b shows an observed morning and afternoon exhibiting the features just discussed.

On almost all mornings, but especially in winter, a small kink occurs in signal strength records such as those of Fig. 4, when the Sun's zenith distance is 95 degrees. On the complex plane the phenomenon appears as a small loop or as a bunching up of the parametric χ marks, as in Figs. 3 and 5. It is not observed in the afternoon. The seasonal variation and absence in the afternoon are not as yet explicable, but the form of the effect produced on the complex plane is what would be expected of a weak ray reflected twice from $D\beta$.

The points in the present sub-section giving strongest support to the two-layer model are the explanations of the seasonal variation of size of the daylight part of the loci, and of the asymmetrical diurnal behaviour in winter.

Aberdeen loci: Loop perpendicular to plane of propagation

Because of experimental difficulties there is only a little observational material relating to the abnormal* component of the field at Aberdeen. Owing to the grazing incidence, a once reflected ray does not produce a strong abnormal component. The polarization of the ray once reflected from $D\beta$ is not known, but so far it cannot be proved that it is not purely normally polarized. The polarization of the multiple reflections from $D\alpha$ is known, and we can calculate the abnormal field to be expected. It is found that in daylight in winter the abnormal component of the ray reflected twice from $D\alpha$ is half the normal component. The rays reflected three and four times are each stronger than that reflected twice, and higher orders are important. In daylight in summer all the rays reflected from $D\alpha$ are negligible. From these numerical results we conclude that the winter loci will be too complicated to analyze easily, and that by sunrise in summer, the nocturnal abnormal field will have shrunk away to nothing. It can only be said at present that the observations so far are not in disagreement with these conclusions.

IV. Relation to Other Work

Two hypothetical regions D_1 and D_2 were found convenient by Bureau and Jouaust [7] in attempting to explain why a sunset dip was noted in records of atmospherics made on 18 kc/sec and higher frequencies, but not on lower frequencies. The dip was considered due to passage of the angle of incidence through the pseudo-Brewster angle, but by admitting a lower layer D_2, reflecting only frequencies lower than 18 kc/sec, and disappearing at sunset, it was considered possible that the Brewster angle phenomenon need not make itself felt at the lower frequencies as reflection jumped from the lower to the higher level. To explain the observed increase in signal strength from day to night on frequencies higher than 18 kc/sec, it was supposed that some absorption of the waves penetrating the lower layer

* The standard terms "abnormal" and "normal" refer to the components of the downcoming wave which have their electric fields respectively perpendicular and parallel to the plane of propagation.

occurred. The two layers were called D_1 and D_2. These surmises do not make immediate contact with the present work, because they are concerned with much greater transmission distances and partly with higher frequencies.

Different explanations of sudden ionospheric disturbances in terms of D_1 and D_2 have been given by JOUAUST [7] and by BUREAU [8], but need not detain us here.

WILKES [9] proposed an ionospheric structure closely resembling the present model. He noted that the diurnal fall in height of 18 km on 16 kc/sec at steep incidence in summer, occurred without any change in reflection coefficient. It seemed improbable that any absorption at the reflecting level could remain unaffected by the great change in collision frequency over such a height interval. WILKES suggested that the absorption on 16 kc/sec at steep incidence in summer be supposed due to a lower layer.

Reports of echoes obtained from below E layer with conventional ionospheric sounding equipment have been summarized by ELLYETT [10]. Many of these reports refer to echoes of a sporadic nature; so far no definite picture of regular ionospheric structure has emerged which could be connected with the present work.

Interesting echo observations on 50 and 100 kc/sec have been made in America [11], [12]. Simultaneous reflections from more than one level have been observed, and we may look forward to results bearing on the present work from experiments of this kind, particularly if they are continued at oblique incidence.

VOL. 2 1952

Finally we may note that a first attempt has been made to measure ionization densities by firing a V 2 rocket into the ionosphere, and that the results appear to indicate a maximum of ionization below E layer [13].

V. DISCUSSION

We have shown that early morning observations of the field of GBR at Aberdeen, when compared with observations at Cambridge, indicate the existence of two physically distinct ionized layers. Eclipse observations by BRACEWELL [14] show that the two layers are distinct from E layer. Neither behaves as a Chapman layer.

Two important earlier surmises regarding the existence of two low-lying layers, both invoke the lower of the two to absorb waves reflected from the higher. We are thus led to consider whether the summer absorption of 16 kc/sec waves, reflected at steep incidence from $D\alpha$, may not take place in $D\beta$. STRAKER [1] has shown that the absorption sets in steadily during morning twilight, when the Sun's zenith distance is decreasing from about 100 to about 90 degrees. In the evening roughly the reverse occurs.

The diurnal variation of height of the reflecting level in $D\beta$ thus agrees with the diurnal variation of absorption. This agreement is very encouraging, because the peculiar diurnal variation of absorption is a serious obstacle to almost any theory. In further support of a connection between $D\beta$ and the absorption of penetrating waves there is the demonstration by STRAKER [1] that the diurnal variation of the amplitude of the waves reflected at steep incidence does not share the asymmetry in the diurnal phase variation of the same waves.

In winter, the twilight fall of the reflecting level in $D\beta$ is practically no different from what is observed in summer. Furthermore, after sunrise the reflecting level in $D\alpha$ is practically the same, for the same value of χ, in summer and winter. And

yet in winter the absorption is much less than in summer. If $D\beta$ is to be the site of the absorption, the difference must lie in the ionization density above the reflecting level in $D\beta$. Reflection at oblique incidence is to be regarded as taking place from the bottom of $D\beta$. Absorption at steep incidence occurs, however, throughout $D\beta$ particularly at the top, since absorption diminishes with collision frequency at these low radio frequencies. We therefore require a layer, strongly ionized in summer and less so in winter, from the bottom of which reflection is strong and practically independent of season. We cannot confirm that this is possible, but nevertheless propose it tentatively.

$D\alpha$ and $D\beta$ may also be expected to attenuate the short waves which pass through them. Appleton and Piggott [15] have shown that the absorption of short waves occurs chiefly below (the maximum of) E layer. Can the absorption be attributed mainly to one or other of our two regions? Short wave absorption is found experimentally to be proportional to $\cos \chi$ where χ is the Sun's zenith distance and n is approximately unity. Such a variation conforms better with the known diurnal height variation of $D\alpha$. If $D\alpha$ is indeed the main short wave absorbing layer, one might expect the diurnal variation of absorption to exhibit the slight asymmetry found in the height variation of $D\alpha$. This matter has been investigated by W. R. Piggott, using the absorption measurements made at Slough. We have been informed privately that there does exist a small degree of asymmetry. Its magnitude is expressed by attributing to the absorbing ionization a relaxation time of 2000 sec. This is equivalent to symmetry about a time 33 min after noon. The corresponding value from the height variation of $D\alpha$ was 25 ± 5 min.

No reference has yet been made to the curious fact that sunlight is able to affect the ionosphere when the Sun is still depressed 8 degrees below the horizon. Visible light from the Sun could be falling on the ionosphere at that time, but such rays would be incident on the ionosphere from below, having passed completely through the ionosphere once. Ionizing constituents would therefore be expected to be absent. Evidently some novel phenomenon is involved. Perhaps visible light is able to liberate electrons under the circumstances, or perhaps some constituent of the atmosphere exists in insufficient quantities to completely use up, on its first passage of the ionosphere, the radiation which ionizes the constituent. This is speculation. It is clear, however, that experiments of the present kind bear closely on attempts which are being made to discover the processes responsible for ionization in the lower ionosphere [16], [17], [18].

VI. Conclusion

We have shown that the observations of GBR at Aberdeen suggest the existence of two ionized layers below E layer. Explanations of many details of the observations have been attempted with reasonable success, even though a rigid separation into pre- and post-sunrise parts was imposed on the explanation for the sake of simplicity.

An anomaly observed by Weekes in the ground interference pattern at distances of about 300 km, has received an explanation in terms of two layers.

The two layer model affects work dealing with the propagation of waves of 16 kc/sec and higher frequencies, including atmospherics, with sudden ionospheric disturbances, and with the absorption of short waves. It profoundly affects theoretical work on the mechanism of reflection of very long waves. The two layer

model also clearly bears on the problem of the location of the electric currents responsible for the solar daily variation of the Earth's magnetic field, and for the magnetic crochet accompanying solar flares. It is confidently expected that the model presented here, though no doubt imperfect, will have a strong influence on future work in these fields.

Acknowledgements—The analysis described in this paper was carried out by R. N. BRACEWELL as part of the programme of radio research of the *C.S.I.R.O. Division of Radiophysics*, Sydney, Australia. The new observational material was obtained in the *Natural Philosophy Department* of *Marischal College, Aberdeen*, by W.C.BAIN, who has continued the oblique incidence experiments begun originally as part of the programme of the *Cavendish Laboratory* by R. N. BRACEWELL and T. W. STRAKER. Our colleague, Dr. STRAKER, shares with us the inference that two layers enter into the propagation of 16 kc/sec waves from Rugby to Aberdeen.

REFERENCES

[1] STRAKER, T. W.; The ionospheric propagation of radio waves of frequency 16 kc/s over short distances (Not yet published). [2] BRACEWELL, R. N.; The ionospheric propagation of radio waves of frequency 16 kc/s over distances of about 200 km (Not yet published). [3] BAIN, W. C., BRACEWELL, R. N., STRAKER, T. W. and WESTCOTT, C. H.; The ionospheric propagation of radio waves of frequency 16 kc/s over distances of about 550 km (Not yet published). [4] WEEKES, K.; Proc. Inst. Elect. Engrs 1950 (Part III) **97** 100. [5] BUDDEN, K. G., RATCLIFFE, J. A. and WILKES, M. V.; Proc. Roy. Soc. A 1939 **171** 188. [6] BRACEWELL, R. N., BUDDEN, K. G., RATCLIFFE, J. A., STRAKER, T. W. and WEEKES, K.; Proc. Inst. Elect. Engrs 1951 (Part III) **98** 221. [7] JOUAUST, R.; L'Ionosphère pp. 179, 217, Paris 1946. [8] BUREAU, R.; L'Onde Electrique 1947, No. 239, p. 3. [9] WILKES, M. V.; Proc. Roy. Soc. A 1940 **175** 143. [10] ELLYETT, C. D.; Terr. Magn. 1947 **52** 1. [11] HELLIWELL, R. A.; Proc. Inst. Radio Engrs 1949 **37** 887. [12] BROWN, J.N. and WATTS, J.M.; J. Geophys. Res. 1950 **55** 179. [13] BERNING, W. W.; J. Meteorology 1951 **8** 175. [14] BRACEWELL, R. N.; Theory of formation of an ionospheric layer below E layer based on eclipse and solar flare effects at 16 kc/sec. J. Atmosph. Terr. Phys. 1952 **2** 226. [15] APPLETON, E.V. and PIGGOTT, W.R.; Proc. Gen. Assy. URSI, Stockholm 1948, p. 320. [16] NICOLET, M.; J. Geophys. Res. 1949 **54** 373. [17] BATES, D. R. and SEATON, S. L.; Proc. Phys. Soc. 1950 **63** 129. [18] PIDDINGTON, J. H.; J. Geophys. Res. The formation of the ionospheric layers (in press).

VOL. 2 1952

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 226 to 235. Pergamon Press Ltd., London

Theory of formation of an ionospheric layer below *E* layer based on eclipse and solar flare effects at 16 kc/sec

R. N. BRACEWELL*

(Received 16 January 1952)

ABSTRACT

The height of reflection of 16 kc/sec waves from the ionosphere at steep incidence was carefully observed during the partial solar eclipse of 28 April 1949, and a symmetrical anomaly was found. This observation, and the observed effects of solar flares, are shown to be incompatible with the theory that reflection takes place from the bottom of a CHAPMAN layer.

An alternative mechanism is proposed whose essential feature is the removal of the restriction in the CHAPMAN theory that the electron density should be negligible compared with the density of ionizable particles. The resulting theory is worked out.

In the extreme case of virtual exhaustion of the supply of ionizable particles in the top of the ionized layer, an increase in the incident ionizing radiation simply lowers the lower boundary of the layer.

The consequences of this sort of behaviour are in agreement with observation as regards (i) the eclipse effect, (ii) the effect of solar flares and (iii) the diurnal variation of height.

INTRODUCTION

It was shown by BUDDEN, RATCLIFFE and WILKES [1] that, if 16 kc/sec waves incident steeply on the ionosphere were assumed to be reflected from the bottom of a CHAPMAN layer, then the form of the diurnal variation of reflection height could be explained. However, there is a reason for thinking that the reflection does not take place from the bottom of a CHAPMAN layer. When a solar flare occurs the reflecting level moves down several kilometres [2], sometimes as much as 15 km [3]. Now when the incident radiation producing a CHAPMAN layer increases in a certain proportion, the rate of production of ions at all levels increases in the same proportion. There is simply intensification of the layer, but no bodily movement. The level where the ionization density is sufficient to reflect very long waves of a given frequency moves downwards, but by a very small amount, because the ionization gradient is very steep at the bottom of a CHAPMAN layer. Downward movements of several kilometres require enormous increases of the ionization which appear impossible on other evidence.

Thus, on the simple view that the extra ionization produced by a flare results from an increase in the radiation responsible for the pre-existing ionization, the hypothesis of a CHAPMAN layer is called into question. However, there is the logical possibility that the flare radiation is not a mere increase of something already present.

A solar eclipse provides an opportunity for testing the matter. CHAPMAN theory indicates, for the physical reason already mentioned, that a decrease in the incident radiation should produce little effect on the height of reflection. In the case of the eclipse to be discussed, the calculated effect would be unobservable.

A small but definite anomaly was, however, observed.

Section 1 of the paper presents the eclipse results and some immediate conclusions. Section 2 develops a new theory of formation of ionized layers to a sufficient extent

* Division of Radiophysics, Commonwealth Scientific and Industrial Research Organization, Australia.

to apply it to the interpretation of experimental results. Sub-section 2.2 includes a discussion of the observed diurnal variation in height of reflection in terms of the new theory, and sub-section 2.3 bears on the question of where short waves are absorbed. Section 3 discusses the eclipse observation and the effect of solar flares in relation to the new theory.

1. The Partial Solar Eclipse of 28 April 1949

1.1. *The expected effect of the eclipse*

The rate of production of ions per unit volume in a Chapman layer is given by the formula (*e.g.* Mitra [4])

$$q = \beta A I_0 P_0 \exp\left\{-\frac{h}{H} - A P_0 H \sec\chi\, e^{-\frac{h}{H}}\right\}$$

where β is the number of ion pairs liberated by unit energy of the incident radiation I_0, ϱ_0 is the density of the atmosphere at ground level, h is the height above the ground, H is the scale height of the (isothermal) atmosphere, A is the absorption coefficient per unit mass, and χ is the Sun's zenith distance. At a height h_1 where the rate of production of ions has the value q_1, we have

$$\frac{h_1}{H} + A P_0 H \sec\chi\, e^{-\frac{h_1}{H}} = \log\frac{\beta A I_0 P_0}{q_1}.$$

Near the maximum of ion production the two terms on the left are of comparable magnitude, but well below the maximum, the second term preponderates. Then the formula for h_1 is

$$h_1 = H\log\sec\chi + H\log A\varrho_0 H - H\log\log\frac{\beta A I_0 P_0}{q_1}.$$

Experiment confirms this formula, inasmuch as h_1 is found to be a linear function of $\log\sec\chi$. The coefficient of $\log\sec\chi$ is found to be 5·5 km for 16 kc/sec waves reflected at steep incidence [5]. Differentiating with respect to I_0, and writing $h(\chi) = H\log A P_0 H\sec\chi$, where $h(\chi)$ is the height of maximum ion production, we have

$$dh_1 = -H\, e^{-\frac{h(\chi)-h_1}{H}}\frac{dI_0}{I_0}.$$

The change in height of a level of constant ion production, initially 2 scale heights below the maximum, produced by a 30% reduction in I_0 (30% of the Sun was obscured during the eclipse) would be

$$dh_1 = -5{\cdot}5\, e^{-2}(-0{\cdot}3) = 0{\cdot}2\text{ km}.$$

This corresponds to a change of phase of 7 degrees, which, spread over the duration of the eclipse, would be on the limit of detectability. Levels more than 2 scale heights down would be affected even less.

1.2. *Details of the eclipse*

The circumstances of the eclipse for Cambridge are given in the Nautical Almanac for 1949. At ground level, the eclipse began at 6 h 22 min, reached its maximum phase at 7 h 17 min and finished at 8 h 16 min. The magnitude of the eclipse is given

VOL. 2 1952

as 0·42, from which the greatest eclipsed area can be calculated to be 0·3. The eclipsed area is plotted against time in Fig. 3. The Sun's zenith distance was 75 degrees at the beginning of the eclipse, and 58 degrees at the end. At mid-eclipse, the heliographic latitude of the centre of the Sun's disk was 4°·42 S, and its heliographic longitude was 236 degrees. In the ionosphere the eclipse occurred approximately one minute earlier, but no special calculations have been made.

Fig. 1 was calculated form the data in the Nautical Almanac. It shows that the obscured area rolled across the northern sunspot zone from west to east. The author is indebted to Monsieur L. D'AZAMBUJA for a fine spectroheliogram, taken in $H\alpha$ at Meudon at the time of the eclipse, which shows three important active regions on the Sun, and four or five others. The position and importance of these regions is indicated in Fig. 1 [6]. Two of the important regions were occulted during the course of the eclipse.

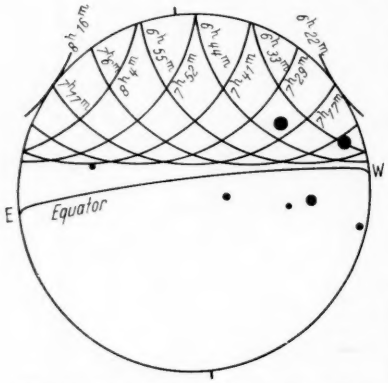


Fig. 1. The Sun's disk, showing the progress of the partial eclipse of 28 April 1949. Active regions are represented by black circles.

1.3. *The observed eclipse anomaly*

Observations were made with the apparatus described by STRAKER [5]. Radio waves which have travelled from the Post Office transmitter GBR Rugby (16 kc/sec) via the ionosphere to Cambridge 90 km away, are compared in phase with the ground wave. The apparatus measures changes in phase caused by changes in height of the reflecting level. A fall in height of three kilometres produces a phase advance of 100 degrees.

Fig. 2 shows the phase measurements made automatically on the control days, 27, 29 April and 2, 3, 4, 5 May 1949, as dots at three minute intervals. The observations exhibit the usual morning fall in height of reflection which is repeated with great regularity from day to day. The time scales of the different days have been displaced slightly to bring the times of sunrise into agreement.

Careful manual observations were taken on the day of the eclipse, and they are shown by the full line. A small but definite anomaly can be seen at the time of the eclipse.

Fig. 3 reveals the form of the anomaly as obtained by substracting an average curve for the control days from the eclipse curve. The obscured area of the Sun's disk has been plotted against time for comparison. A zero error of about 10 degrees which appears to be present is due to the experimental uncertainty in determining the phase of the ground wave on the eclipse day. Fluctuations amounting to about 10 degrees are normal, and are of ionospheric origin.

The size of the anomaly is estimated to be 35 ± 5 degrees at mid-eclipse, or 1·0 km. It may also be seen to be to a first approximation symmetrical, and to start and finish at about the times of the optical eclipse of the visible solar disk.

Careful study of the data and experience of this type of ionospheric record indicate that the conclusions of the preceding paragraph are the most definite warranted by the observations. However to give them a more quantitative aspect

we note that the maximum and start of the eclipse anomaly appear to agree with the visible eclipse within about 10 min, this figure being admittedly rough. The end of the anomaly came later than the end of the eclipse, but not, it is considered, significantly later.

1.4. *Deductions from the observations*

The size of the anomaly supports the theory that reflection does not take place from a CHAPMAN layer, because it was several times bigger than expected. This was the main object of the work, but a number of further deductions are suggested.

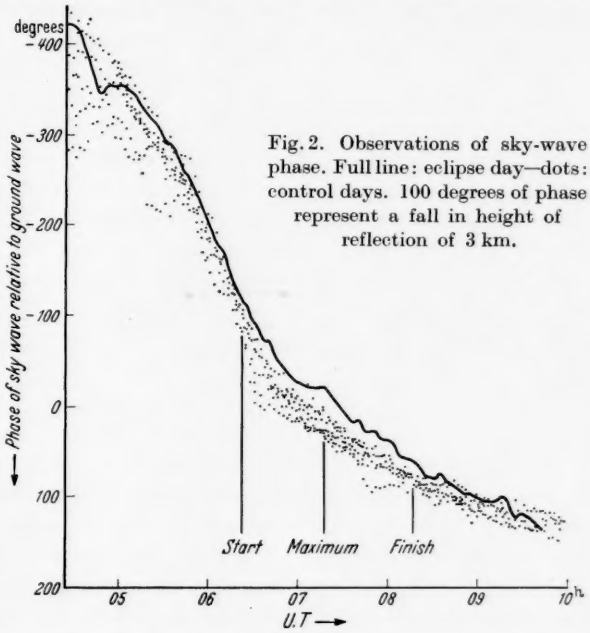


Fig. 2. Observations of sky-wave phase. Full line: eclipse day—dots: control days. 100 degrees of phase represent a fall in height of reflection of 3 km.

Fig. 2.

A certain type of conclusion regarding the level of origin of the ionizing radiation in the Sun's atmosphere can be drawn from eclipse anomalies. Thus if we find that the anomaly began no earlier than 10 min before first contact, we can calculate that the ionizing radiation does not originate in the corona at a height greater than 170,000 km above the photosphere. Tihs is too high to be of much interest.

The form of the anomaly can be seen to agree roughly with the shape of the obscured area curve. If Fig. 1 is studied in conjunction with Fig. 3 it will be found that there is no correspondence between the fluctuations in the record and the occulting of the active regions. Furthermore the fluctuations are of the size usual at all times. Consequently it may be deduced that, within the experimental accuracy the ionizing radiation originates regularly from all over the Sun. It does not originate to an important extent in active regions.

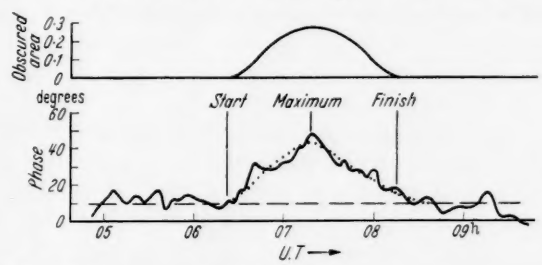


Fig. 3. Eclipse anomaly obtained by subtracting mean of control days from eclipse day.

It is possible that the radiation is not uniformly distributed but shows a centre-limb variation. The form of the anomaly is not very sensitive to the precise law of darkening, but does in principle yield it, provided a theoretical relation between amount of radiation and amount of anomaly is available. This is discussed later.

VOL. 2 1952

From the symmetry of the anomaly we deduce a short relaxation time for the ionization in the layer. Fig. 3 indicates that it can hardly exceed 6 min. On the other hand a degree of asymmetry in the diurnal variation of height has been found by BRACEWELL [7], which, interpreted as a recombination phenomenon, gives a relaxation time of 25 min. The present observation is, however, in agreement with the observations of BRACEWELL and STRAKER [2], [8], who showed from the rapidity with which sudden phase anomalies due to solar flares recover, that the relaxation time constant does not exceed 5 min and may be much less. The ability of the layer to respond rapidly to short term variations in radiation is thus established both for increases and decreases. The diurnal asymmetry now appears not to be simply a recombination phenomenon, and will be discussed elsewhere.

2. THEORY OF EXHAUSTION REGIONS

Observation shows that a decrease in I_0 raises the bottom of the reflecting layer. Conversely an increase lowers the reflecting level. The height changes are much greater than CHAPMAN theory indicates. We seek some reason why a moderate increase in I_0 should not simply intensify the existing layer, but instead expend itself producing new ionization below that already existing. The existing ionization must modify the capacity of the atmosphere to be ionized. This is explicitly assumed not to be so in the CHAPMAN theory where electrons are supposed to be much less numerous than ionizable particles. This assumption must be abandoned in the face of the observations. The simplest way in which the production of ions may reduce the capacity for further ionization is by depletion of the supply of ionizable particles. We therefore adopt the hypothesis that the atmospheric constituent from which electrons are knocked by the ionizing radiation, exists in a small concentration, and is ionized to a not negligible degree. The theory of ionized layers, with the usual assumption of negligible degree of ionization removed, now follows.

2.1. *Basic theory*

Consider an isothermal atmosphere with scale height H in which the density of ionizable particles n varies exponentially with the height h as given by

$$n = n_0 e^{-\frac{h}{H}}.$$

Let I be the intensity of vertically incident monochromatic radiation, and I_0 its value outside the atmosphere. Let the absorption coefficient per particle be A, and let β electrons be produced per unit of energy absorbed. Let q be the rate of production of electrons per unit volume, and n_e the electron density.

Then

$$\frac{dI}{dh} = AI(n - n_e) \tag{1}$$

and

$$q = \beta \frac{dI}{dh}. \tag{2}$$

Let the rate of disappearance of electrons per unit volume be αn_e^m so that in equilibrium

$$q = \alpha n_e^m.$$

To solve these equations for the variation of n_e with h we must adopt a specific law for the disappearance of electrons: let $m = 1$. Eliminating q we have the differential equation for I,

$$\frac{dI}{dh} = \frac{n_0 e^{-\frac{h}{H}}}{\frac{\beta}{\alpha} + \frac{1}{AI}}.$$

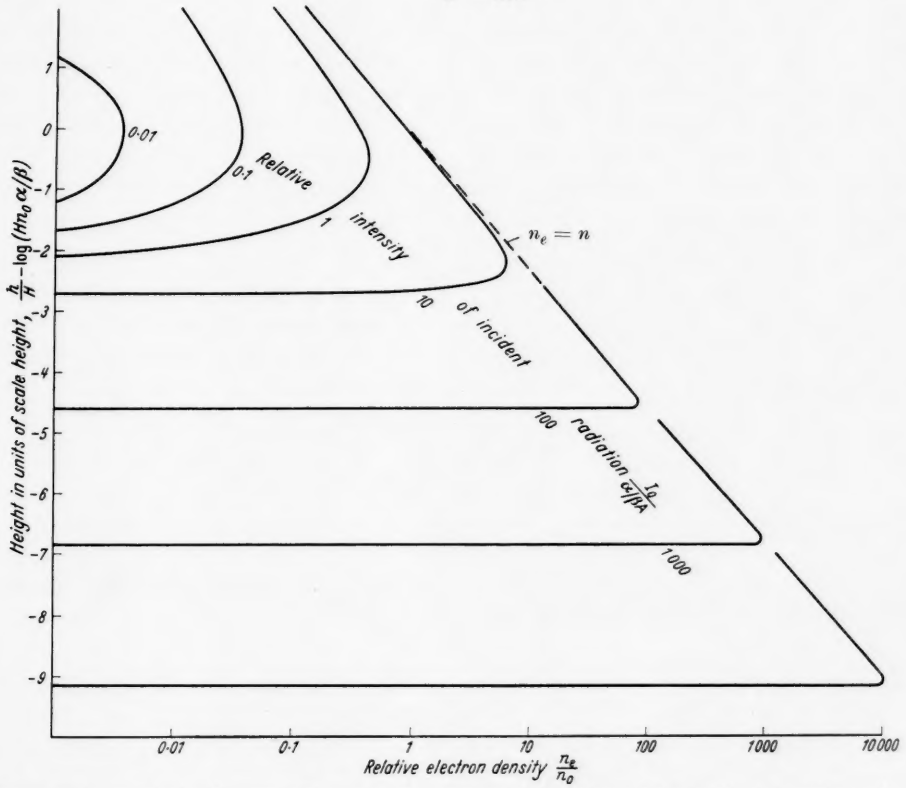


Fig. 4. Universal ionization—height curves for different intensities of the incident ionizing radiation I_0. $I_0 \ll \alpha/\beta A$ corresponds to CHAPMAN layers and $I_0 \gg \alpha/\beta A$ to "exhaustion regions."

Integrating,

$$I - I_0 + \frac{\alpha}{\beta A} \log \frac{I}{I_0} = - H n_0 \frac{\alpha}{\beta} e^{-\frac{h}{H}}.$$

This gives the variation of I with h. A relation connecting n_e with I is then deduced by eliminating dI/dh between [1] and (2). Thus

$$n_e = \frac{n_0 e^{-\frac{h}{H}}}{1 + \frac{\alpha}{\beta A I}}.$$

From these last two equations one can construct universal curves representing the distribution of n_e with h for different values of I_0. These are shown in Fig. 4.

VOL. 2 1952

A graphical representation covering all cases by means of a single family of curves is made possible by use of the dimensionless variables n_e/n_0, h/H and $I_0(\alpha/\beta A)^{-1}$. The term $\log(H n_0 \alpha/\beta)$, which depends on the properties of the atmosphere, allows for layers of the same form situated at different heights, by a shift of origin.

As the present theory includes the CHAPMAN theory as a special case, the familiar properties of the latter are recoverable under the condition that I_0 should be small compared with $\alpha/\beta A$. Thus in Fig. 4 we see that when $I_0(\alpha/\beta A)^{-1}$ is small, then n_e is small compared with n, and an increase in I_0 intensifies the existing ionization without much effect on the form of the region or its height of maximum ionization.

On the other extreme, when I_0 is large compared with $\alpha/\beta A$, we find $n_e \doteqdot n$ down to a certain level, below which n_e falls rapidly to small values. An increase in I_0 then produces new ionization practically all below the existing ionization. When I_0 has intermediate value, *i.e.* when $I_0(\alpha/\beta A)^{-1} \doteqdot 1$, an increase in I_0 both intensifies and lowers the existing layer.

Having now given a quantitative illustration of the effect of relieving the assumption that $n_e \ll n$, we shall confine ourselves to situations where I_0 is strong. We now have a simple physical situation where the ionizable particles are practically fully ionized down to a certain level, and practically un-ionized below. This is what is referred to as an exhaustion region. The connection between I_0 and the height of this level, h_0, is given by the condition

$$I_0 = \int_{h_0}^{\infty} \frac{dI}{dh}\, dh$$

whence

$$I_0 = \frac{\alpha}{\beta} n_0^m \frac{H}{m} e^{-\frac{m h_0}{H}}.$$

Suppose that I_0 changes to I_0' and that h_0 changes correspondingly to h_0'. Then for the new height h_0' we have finally the convenient formula

$$h_0' = h_0 - \frac{H}{m} \log \frac{I_0'}{I_0}.$$

It will be noted that this formula is not restricted to $m = 1$, as in the general analysis carried through for the purpose of Fig. 4. Some implications of the formula are now derived and compared with experience.

2.2. *Diurnal variation of height*

If the incident radiation comes from a direction inclined at an angle χ to the vertical it penetrates to a height h_0' which can be calculated by taking $I_0' = I_0 \cos\chi$. Thus

$$h_0' = h_0 + \frac{H}{m} \log \sec\chi.$$

The expression for the observed diurnal variation is

$$5{\cdot}5 \log \sec\chi + \text{const}.$$

The theory therefore correctly predicts the form of the diurnal variation of height, and gives the experimental value 5·5 km for H/m.

It may be recalled from Section 1 that CHAPMAN theory leads to the same form of diurnal variation. The coefficient of $\log \sec\chi$ there came out to be H, irrespective of the value of m. The satisfactory agreement between theory and observation led

BUDDEN, RATCLIFFE and WILKES [1] to consider that they had measured H, and they deduced from it a widely-quoted value for the temperature. According to the present theory, however, we do not know H, m or T, but only the ratio H/m. In recent papers measurements of the "height excursion coefficient" are no longer considered to be convertible into temperature, nor are seasonal changes in the coefficient now thought to necessarily imply seasonal temperature changes.

2.3. *Absorption of short waves*

It is of interest to calculate the absorption suffered by short waves penetrating an exhaustion region. Let the collision frequency ν vary exponentially with height, $\nu = \nu_0 \exp(-h/H)$, and let the electron density be given by $n_e = n_0 \exp(-h/H)$ above h_0' and zero below. The absorption coefficient of the ionized medium for the ordinary wave is $\frac{2\pi e^2}{m c} n_e \nu [\nu^2 + (p + p_L)^2]^{-1}$, where p is the angular frequency of the wave and p_L the longitudinal gyrofrequency, so provided that $\nu^2 \ll (p + p_L)^2$, the total absorption is proportional to the integral of $n_e \nu$ over the whole region. Thus

$$\int_{h_0'}^{\infty} n_e \nu \, dh = n_0 \nu_0 \int_{h_0'}^{\infty} e^{-\frac{2h}{H}} dh = n_0 \nu_0 \frac{H}{2} e^{-\frac{2h_0'}{H}}.$$

VOL. 2 1952

Now h_0' varies through the day according to the equation

$$h_0' = \frac{H}{m} \log c \sec \chi$$

where c is a constant;

$$\therefore \; n_0 \nu_0 \frac{H}{2} e^{-\frac{2h_0'}{H}} = n_0 \nu_0 \frac{H}{2} c^{-\frac{2}{m}} \cos^{\frac{2}{m}} \chi.$$

For $m = 2$ the diurnal variation is thus proportional to $\cos \chi$ and for $m = 1$ it is proportional to $\cos^2 \chi$.

On CHAPMAN theory the exponent of $\cos \chi$ is $\frac{3}{2}$. Experimental values range from 0·7 to 1·3 [9].

It is considered probable that this region is responsible for a great part of the absorption of short waves. Thus APPLETON and PIGGOTT [10] have shown that the absorption occurs below E layer. BRACEWELL and BAIN [11] have pointed out that short wave absorption shows the same diurnal asymmetry as one of the two layers which they distinguish below E layer, *viz.* the region which reflects 16 kc/sec waves at steep incidence. Finally we find that the calculated diurnal variation of short wave absorption, assumed due to an exhaustion region, agrees reasonably with observation.

3. DISCUSSION IN RELATION TO THE NEW THEORY

From the formula

$$h_0' - h_0 = -\frac{H}{m} \log \frac{I_0'}{I_0}$$

the magnitude of the eclipse effect can be predicted if the coefficient H/m is known. A rough direct estimate is possible. The temperature of the atmosphere at the height involved, *viz.* between 60 and 80 km is probably between 200 and 350° K (*e.g.* BATES [12]). The scale height, calculated from the expression kT/mg is thus from

6 to 10 km. The value of m is expected to be about 1 or 2. Consequently H/m is of the order of magnitude of 5 km. We shall, however, use the experimental value of 5·5 km for H/m, based on the height changes observed at the obliquity of the Sun's rays changes (sub-section 2.2). Remembering that the unobscured proportion of the Sun's disk as mid-eclipse was 0·7, we have $I_0'/I_0 = 0{\cdot}7$ and $H/m = 5{\cdot}5$ km, whence the calculated anomaly is $5{\cdot}5 \log 0{\cdot}7 = -2{\cdot}0$ km. This is a bigger effect than was observed, but it is necessary to take account of possible limb-darkening. A reasonable extrapolation of the observed coefficients of limb-darkening (ABBOTT [13]) is to take a value of unity for ultra-violet light. This means a simple cosine distribution over the disk. The partial eclipse then reduces the radiation by a lesser extent than before. We find $I_0'/I_0 = 0{\cdot}86$, and the calculated anomaly becomes −0·8 km, compared with the observed value 1·0 km. This is good agreement with theory.

The form of the eclipse anomaly is insensitive to the precise law of limb-darkening, and was not well enough determined in the present case to give additional information.

Sudden phase anomalies due to solar flares show just the reverse behaviour to that seen during the eclipse. According to the theory, if the incident radiation shows a transient increase by a factor $A(t)$, which represents the course of a flare, then a negative height anomaly is observed of the form $-5{\cdot}5 \log A(t)$ km. Thus a large anomaly with a maximum of 15 km would be produced by an increase of $\exp(15/5{\cdot}5) = 15$ times in the incident radiation. It is of interest that the theory leads to a geophysical measure of the relative increase in ultraviolet radiation at the time of a solar flare. This figure will be useful for comparison should a theoretical or observational estimate become available from another source.

To illustrate the enormous increases in ionization implied by CHAPMAN theory, we may take as an example a large sudden phase anomaly involving a fall of two scale heights from a level initially two scale heights below the maximum. The increase in maximum ionization density would be 10^{10} times, and the critical frequency would go up 10^5 times. There would be similar increases in absorption. On the new theory, observed anomalies imply moderate increases more in keeping with measurements of critical frequency and estimates of total absorption made during sudden ionospheric disturbances.

As long as it could be considered that 16 kc/sec waves were reflected from the bottom of a CHAPMAN layer, it was possible that reflection took place from the bottom of E layer, which behaves reasonably in agreement with CHAPMAN theory. The height of reflection is not known accurately enough to rule out this possibility. The eclipse result shows, however, that we are not dealing with a CHAPMAN layer. It follows that there exists below the E layer a regular ionized layer which is physically distinct from it.

It is believed that this is one of the first pieces of clear-cut evidence for this conclusion. It might have been thought, from the frequency with which the term D region appears in publication, that it denoted some single physical entity with generally accepted properties, but that is not the case. The author has introduced the notation $D\alpha$ for the layer under discussion to avoid the confusion associated with the term D region, and to distinguish it from a second layer also below E layer [11].

4. Conclusion

A solar eclipse produces an observable anomaly in the height of reflection of 16 kc/sec waves at steep incidence. The anomaly is regular and symmetrical and agrees in duration with the optical eclipse. The ionizing radiation originates regularly over the Sun's disk and not, to an important extent, in active regions associated with sunspots, nor in the high corona.

A former theory of the ionized layer is not capable of explaining the eclipse effect or the effect of solar flares. An alternative theory has been developed which predicts an eclipse anomaly. This theory accounts for sudden phase anomalies due to solar flares in terms of moderate increases of radiation, and offers an absolute measure of intensity of solar flares.

The present theory has been presented in the simplest form embodying the essential idea of exhaustion of the supply of electrons, so that the consequences peculiar to such a circumstance could be illustrated free from complications. It is expected that a more accurate theory will have to take account of the variation of temperature with height, the probable variation of the effective recombination coefficient with pressure, and the distribution of the ionizable material. Meanwhile we have a broad working hypothesis asserting the existence of an ionized layer and the physical principle by which it is formed. Even the elementary facts about this region are still unknown; *e.g.* what the ionized substance is. The present tentative picture is of a flood of ultraviolet radiation, too long in wavelength to ionize the upper ionosphere, streaming down through *F* and *E* layers and encountering a concentration of some easily ionized atmospheric constituent such as sodium, or some negative ion, which it fully ionizes down to the level where it is finally all absorbed.

This first observation of a long wave eclipse effect on the ionosphere has proved fruitful in stimulating a new view on the little-known part of the ionosphere below *E* layer. It has provided factual material for use in theories of the formation of the ionized layers, and in theories of solar flares, both difficult fields in which new facts are likely to be welcome.

Acknowledgements—The work described was carried out partly in the Radio Section of the *Cavendish Laboratory*, Cambridge, during the tenure of a C.S.I.R.O. Studentship, and was completed at *Radiophysics Laboratory*, Sydney. The Post Office kindly provided some special transmissions from GBR Rugby. The eclipse observations were planned in collaboration with Dr. T. W. Straker, who took the records.

References

[1] Budden, K. G., Ratcliffe, J. A. and Wilkes, M. V.; Proc. Roy. Soc. A 1939 **171** 188. [2] Bracewell, R. N. and Straker, T. W.; M.N.R.A.S. 1949 **109** 28–45. [3] Straker, T. W. and Bracewell, R. N.; The Observatory 1949 **69** 69–70. [4] Mitra, S. K.; The upper atmosphere. Calcutta 1948. [5] Straker, T. W.; The ionospheric propagation of radio waves of frequency 16 kc/sec over short distances (not yet published). [6] Quarterly Bulletin on Solar Activity No. 86, Zürich 1949. [7] Bracewell, R. N.; The ionospheric propagation of radio waves of frequency 16 kc/sec over distances of about 200 km (not yet published). [8] Bracewell, R. N. and Straker, T. W.; The Observatory 1949 **69** 143–145. [9] Taylor, E.; J. Research Nat. Bureau Stand. 1948 **41** 575. [10] Appleton, E. V. and Piggott, W. R.; Proc. Gen. Assy. URSI, Stockholm 1948, p. 320. [11] Bracewell, R. N. and Bain, W. C.; An explanation of radio propagation at 16 kc/sec in terms of two layers below *E* layer. J. Atmosph. Terr. Phys. 1952 **2** 216. [12] Bates, D. R.; M.N.R.A.S. 1949 **109** 215–245. [13] Abbott, C. G.; Smithsonian Annals 1922 **4** 229.

VOL. 2 1952

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 236 to 252. Pergamon Press Ltd., London

The morphology of storms in the F_2 layer of the ionosphere

I. Some statistical relationships

E. V. Appleton F.R.S.* and W. R. Piggott B.Sc.**

(Received 26 January 1952)

ABSTRACT

The results of a statistical study of ionosphere storm phenomena at a number of stations are described. These have revealed a marked diurnal control of storm manifestation accountable only in terms of station local time. The different average sequences of storm events at auroral, temperate and equatorial latitudes are contrasted, and it is shown that, in temperate latitudes at least, such sequences display a seasonal variation.

1. Introduction

Shortly after the main characteristics of the diurnal and seasonal changes in the ionosphere had been identified it was discovered that large departures from normality occasionally occurred in the F region of the ionosphere. The most striking type of deviation noticed was a large decrease in the maximum ionisation density of the F_2 layer accompanied by a large increase in its apparent height during the day; and abnormal turbulence, height and density at night. It was quickly recognised that many of these disturbances were associated with periods of abnormal magnetic fluctuation [1], but attempts to establish a detailed correspondence between the two groups of phenomena failed.

Observations made at Tromsø (Lat. 70° N) near the auroral belt, during the second International Polar Year 1932–1933 [2] showed striking examples of such F_2 Layer phenomena, which were also often accompanied by large changes in the lower ionospheric layers. For this station the connection between ionospheric and magnetic disturbances was particularly close. It was found, for example, that, under conditions of weak magnetic activity, there often occurred abnormal E-Layer ionisation and reduced maximum F_2 Layer electron density. Under conditions of strong magnetic activity, however, low-level absorption was found to be so intense that, temporarily, no ionospheric echoes at all were detectable with the particular radio equipment then employed. Comparing the results obtained at Tromsø with those at a lower latitude station, Slough ($51\frac{1}{2}$° N), it was found that absorption effects were not well marked at the latter station, the chief ionospheric manifestation being the reduction in the maximum electron density of the F_2 Layer.

Later observations made at many high and low latitude stations have confirmed these early results and, in addition, have established that the great severity of ionospheric disturbances and their close association with magnetic phenomena are general features of high latitude stations.

The most widely recognised statistical correlation between ionospheric and magnetic events is the depression of the F_2 Layer critical frequency (f_0F_2) with increased magnetic activity. This was noted by Appleton and Ingram [1] using Slough measurements of f_0F_2 and Greenwich Magnetic Character Figures. For

* The University, Edinburgh.

** Radio Research Station (D.S.I.R.), Slough, Bucks.

noon conditions it was found that intense magnetic storms were often accompanied by a reduction in f_0F_2, though the same authors added this significant remark "But we have noted a number of cases where the critical frequencies are actually increased when a large storm is in progress." For midnight conditions it was found that, while small magnetic activity could be associated with an increase of f_0F_2, severe magnetic disturbance was associated with a depression of the same quantity. Similar studies were made by HARANG [3], at Tromsø, who found that a more marked inverse correlation was found between magnetic activity and the F_2 Layer critical frequency if the value of the latter was taken for the day subsequent to that selected for the particular magnetic character figure. In other words there was a tendency for the major depression of f_0F_2 to be experienced the day after the magnetically stormy day.

Further statistical studies of the same relation between F_2 Layer phenomena and magnetic character have been made by BERKNER and SEATON [4] using the ionospheric data from Huancayo (Lat. 12° S) and Watheroo (Lat. 30° S). At Huancayo, which is situated on the magnetic equator, the value of f_0F_2 was found to increase steadily with increase of magnetic activity, whatever the season. At Watheroo, on the other hand, there were found to be different results in summer and winter. In summer the average value of f_0F_2 dropped steadily as the value of the magnetic character figure increased. In winter, on the other hand, the average value of the F_2 Layer critical frequency first increased with increase of magnetic activity, but tended to decrease during periods of marked disturbance. APPLETON, NAISMITH and INGRAM [2] had also found a seasonal variation in the correlation between ionospheric and magnetic disturbances at Tromsø for 1932–1933, the correlation between the depression of F_2 Layer critical frequency and magnetic disturbance being more marked in summer than in winter.

VOL. 2 1952

Thus the published data indicate that the correlation between F_2 critical frequency and magnetic disturbance can be different in magnitude and sign at different stations and may vary seasonally. The same data illustrate the fact that the F_2 critical frequency may be either above or below normal during any individual magnetic disturbance at a given site. It is doubtless for these reasons that many workers have doubted whether such correlations as have been mentioned have any fundamental significance.

In some recent work we have attempted to identify some of the chief features of ionospheric storms with the object of gaining insight into the mechanisms which are responsible for the perturbations. For this purpose our analysis has, in the main, involved two complementary techniques;

(a) A statistical investigation for elucidating the features which are common to a large number of storms at either one or more stations.

(b) A detailed study of the sequences of events in selected individual storms.

In Part I of this paper we deal with the results of (a) and, in Part II, with the results of (b).

2. VARIATIONS IN THE F_2 LAYER

Changes in the structure of any ionised layer may be disclosed as changes in (a) maximum electron density, (b) layer height and shape, and (c) degree of turbulence. Such changes are not necessarily similar or simultaneous, and a detailed knowledge

of all three is necessary before a complete description of any particular perturbation can be given. Unfortunately it is only possible to obtain comparable values of the critical frequencies from most stations, the published data on heights and turbulence being, in general, unsatisfactory for our purpose. Most of the discussion below will therefore be confined to an examination of the variation of the maximum ionisation density during ionospheric storms, both with time and with site.

It is very desirable to find a simple parameter, the fluctuations in which can represent the day-time or night-time ionisation density changes from day to day or from station to station. A careful examination of the short period fluctuations in the F_2 critical frequency shows that single observations may be unrepresentative of mean conditions. A suitable index of f_0F_2 may, however, be obtained by averaging not less than five observations separated by intervals of at least a quarter of an hour. For the purpose of examining day-to-day or station-to-station changes in the ionosphere we have therefore used the mean of the five hourly values centred in local noon for the daytime index; and the mean of the five hourly values centred on midnight for the corresponding night-time quantity.

These quantities have been found to be reliable indices of average day-time or night-time conditions, but are obviously inadequate for the study of short lived phenomena.

Before proceeding to a detailed examination of phenomena during magnetic or ionospheric storms it is necessary to examine the data from widely separated stations in order to detect whether there are large world-wide fluctuations in maximum ionisation density which might arise from other causes *e.g.*:—fluctuations in the intensity of the ionising radiation. As is well-known, large long-term changes in ionising radiation occur during the sunspot cycle and since solar activity is often localised, rapid changes in solar ultra-violet light might be expected.

Comparing the deviations of the noon critical frequency index for each day from its monthly running median at widely separated stations we find that the correlation between these two quantities at pairs of stations is zero. Two typical examples illustrating these results are shown in Tables 1 and 2.

We may conclude, therefore, that effects due to day-to-day changes in the ionosing radiation are usually small compared with other perturbations.

A single exception to this conclusion has, however, been noted during the last sunspot cycle. This occurred during March 1942, at a period of sunspot minimum,

Table 1. Comparison of F_2 Layer noon deviations at Johannesburg (26° S) and Slough (51° N), 1948 and 1949. (Same longitude and all days.)

OBSERVED

Johannesburg \ Slough	+	−
+	187	178
−	188	161

Total 714

EXPECTED

Johannesburg \ Slough	+	−
+	192	173
−	183	166

Total 714

(Days with deviation above normal counted + ; below normal −.)

Table 2. Comparison of F_2 Layer noon deviations at Canberra (35° S, 149° E) and Washington (39° N, 77° W), 1943. (Different longitudes and days of large deviation only.)

OBSERVED

Canberra \ Washington	+	−
+	35	38
−	42	54

Total 169

EXPECTED

Canberra \ Washington	+	−
+	35	38
−	45	51

Total 169

(Days with deviation greater than 1 Mc/s above normal counted +; greater than 1 Mc/s below normal −.)

when there was a single highly active spot on an otherwise quiet sun. All three main ionospheric layers exhibited simultaneous increases in ionisation at all stations while the spot was visible and active.

3. The Sequence of Events During Ionospheric Storms

(a) *Day-to-day variations of f_0F_2*

As pointed out above, the most striking feature of an ionospheric storm in temperate latitudes is the depression of the F_2 critical frequency. It is natural, therefore, to examine a number of occasions when this phenomenon has been experienced in order to discover whether it displays any characteristic variations with time.

We have taken the values of the largest departures, Δf_0F_2, from monthly mean Slough values at noon for 40 storms during the period 1942–1944, near Sunspot Minimum, and have set out, by the superposed-epoch method, the average values of Δf_0F_2 for the days of largest departures (zero days) and for days ±1, ±2, ±3, etc. relative to this. The results are exhibited in Fig. 1. Here it is seen that, for two days before the major departure from the monthly mean value, the critical frequency for the F_2 layer was significantly above the average. After the major depression the recovery is seen to be slow.

A similar curve (see Fig. 2) for 34 storms occurring in the early part of the decline of the sunspot cycle (1939–1940) has a somewhat similar general trend, though it is noticed that (a) the depression of f_0F_2 on zero day is greater, (b) the enhanced value of f_0F_2 for the period preceding zero day is not of so long a duration and (c) the recovery to normality is quicker.

These two curves taken together give the starting point of our enquiry, for they show that, associated with an ionospheric storm, there is a positive phase when the critical frequencies are above normal as well as the well known negative phase; and that the positive phase precedes the negative phase in time. Curves of the same general trend have been obtained for other stations at temperate latitudes in both the North and the South hemispheres, the differences (*e.g.* in amplitude and time scales) between the curves at different stations varying slowly with latitude and longitude. If, however, we attempt to use the same procedure for a station near the equator we find that there are very few occasions when there is a sudden drop in F_2 critical frequency. Coherent results can be obtained by selecting as the zero days for the

VOL. 2 1952

equatorial station, days of marked f_0F_2 depression at a temperate latitude station on the same longitude. It is then found that the equatorial station exhibits a quite well marked increase in f_0F_2 roughly simultaneously with the temperate zone depression. This difference in ionospheric storm behaviour in low and temperate latitudes is illustrated in the graphs of Figs. 3 and 4.

The application of the superposed-epoch technique to noon data from auroral stations at high latitudes encounters considerable difficulty since the ionisation

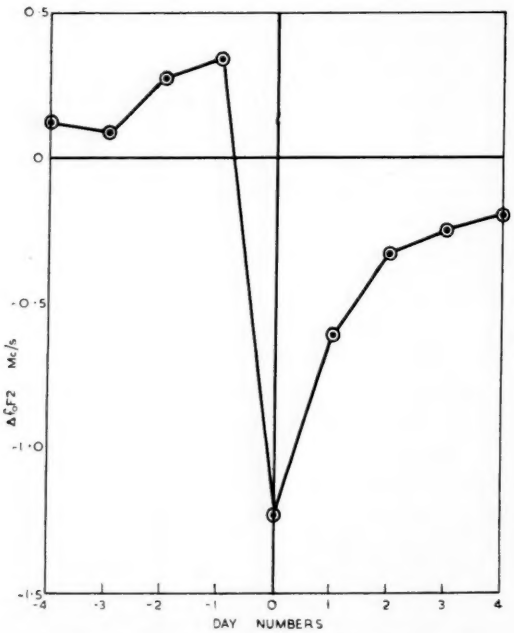


Fig. 1. Superposed epoch diagram for ionospheric storms at Radio Research Station, Slough, near minimum of sunspot cycle.

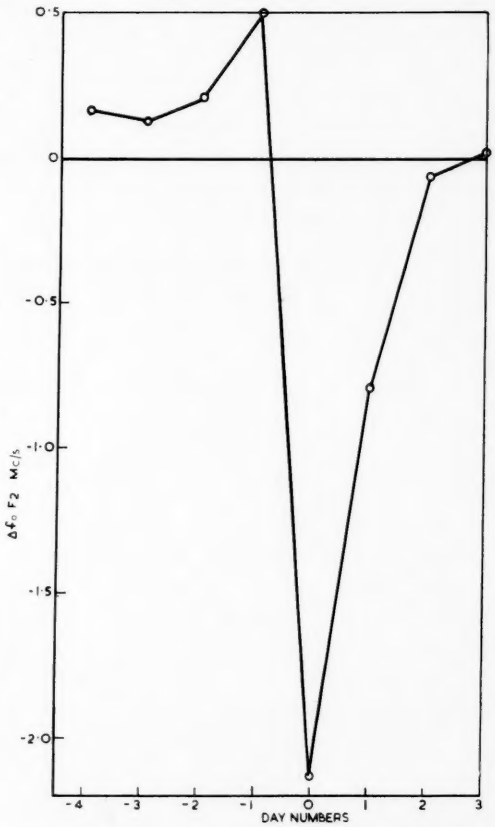


Fig. 2. Superposed epoch diagram for ionospheric storms at Radio Research Station, Slough, near maximum of sunspot cycle.

density there varies rapidly with time. Moreover, for a large part of disturbed periods, screening effects of intense auroral E Layer ionisation and heavy D region absorption are operative. In general, however, we may note that the positive phase is reduced in duration and is often absent and the negative phase is shorter than at lower latitudes.

Having thus established that systematic day-to-day variations occur at many stations outside the auroral zone it is important to discover whether the changes at different stations are associated. The deviations of our five noon-hour indices at each station are very convenient for such analyses and yield results which readily lend themselves to statistical tests of significance.

(b) *Variations of f_0F_2 along a line of longitude*

We first consider the correlation of large positive or negative f_0F_2 deviations at stations sited in a zone of approximately constant longitude. For this analysis we have selected days for which the deviations of the noon index is greater than ± 1 Mc/s at one station, and, using a 3×3 square, have examined the deviations at other stations for these days. We have then selected the days on which the latter deviations have also exceeded ± 1 Mc/s. An example of the results obtained is set out in Table 3 for pairs

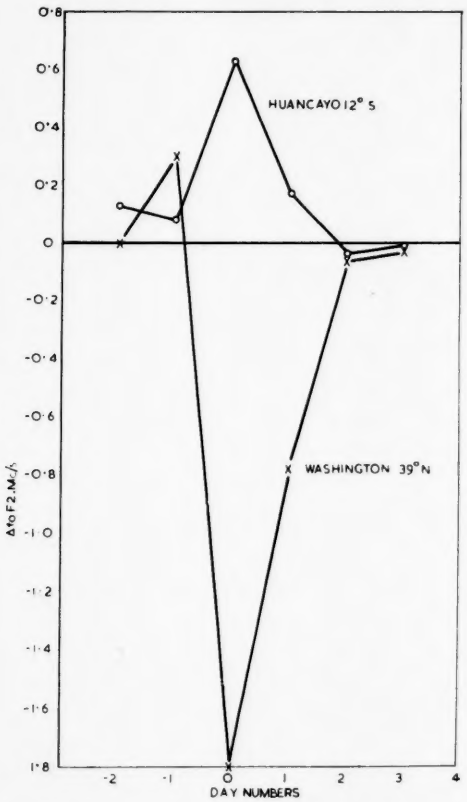


Fig. 3. Superposed epoch curves for Huancayo and Washington. Based on first day when Washington showed a storm effect 1944–1946.

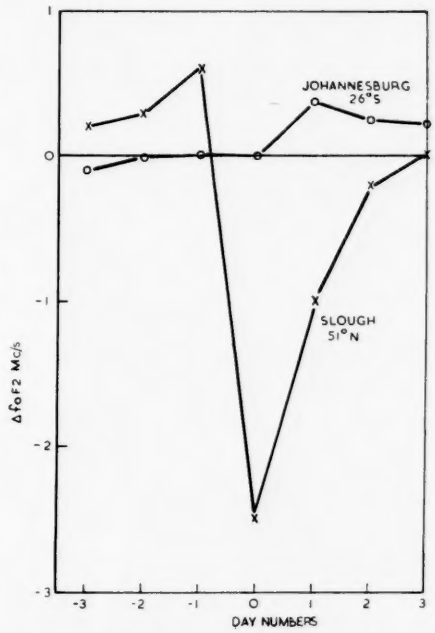


Fig. 4. Superposed epoch curves of Δf_0F_2 at Slough and Johannesburg 1948 using Slough disturbances for zero dates.

of the three stations, Churchill, (Lat. 59° N; Long. 94° W), Ottawa (Lat. 45° N; Long. 76° W), and Washington (Lat. 39° N; Long. 77° W).

It is obvious that conditions at all three pairs of stations are very closely associated. The degree of association decreases as we move to zones of higher activity, although the number and magnitude of the disturbances increase.

If, instead of using large deviations of f_0F_2, we employ all the deviations at a master station, we find that different results are obtained depending on whether the master station is more or less subject to storms than the comparison station. This effect is particularly marked near and in the auroral zone. We may summarise these results by the statement that the correlation is much greater when we select

Table 3. Comparison of F_2 Layer noon deviations at Washington, Ottawa and Churchill, 1948 and 1949. (Similar longitude, and days of large deviations.)

OBSERVED

Ottawa

Washington	+	−	
+	30	0	Total 109
−	0	79	

Churchill

Ottawa	+	−	
+	26	5	Total 88
−	3	54	

Churchill

Washington	+	−	
+	16	4	Total 76
−	5	51	

EXPECTED

Ottawa

Washington	+	−	
+	7	11	Total 49
−	12	19	

Churchill

Ottawa	+	−	
+	10	12	Total 59
−	17	20	

Churchill

Washington	+	−	
+	9	11	Total 58
−	17	21	

(Days with deviation greater than 1 Mc/s above normal counted +; greater than 1 Mc/s below normal −.)

the lesser disturbed station as the master. Thus many disturbances disappear or change in form as we progress from the auroral zone towards the equator.

It is interesting to examine the 1948–1949 results at Churchill for those periods during which both Washington and Ottawa exhibited similar and large deviations. These are shown in Table 4 and show a close association.

Table 4. Churchill deviations

	Mc/s	$>+1$	$+1>0$	$0>-1$	<-1
Ottawa	+	9	16	4	0
Washington . .	−	1	3	26	41

It should be noted that Ottawa and Churchill and, to a lesser extent, Washington, are auroral zone stations showing many of the typical characteristics of Tromsø.

Repeating, for other pairs of stations, the type of analysis which yielded the results set out in Table 3 we find, for stations in temperate latitudes, excellent correlations for the larger deviations at stations on or near the same longitude. For example, the data from Burghead, Slough, Lindau and Graz show correlations comparable with those set out in Table 3. At the lower latitudes the number of

large deviations decreases considerably, as does also the correspondence between pairs of stations.

As we have seen from the superposed-epoch curves (*e.g.* Figs 3 and 4), the storm manifestation at the equator is predominantly an increase in noon ionisation density, such increase being often late relative to the major characteristic decrease found at higher latitudes. In Table 5 the correlation of the Huancayo deviations with periods when the deviations at Washington were large during 1944–1946 and 1948–1949 is shown both for the same day and with the Huancayo data for the day after the Washington effect.

Table 5. Comparison of F_2 Layer noon deviations at Washington and Huancayo, 1944, 1945, 1946, 1948 and 1949.

OBSERVED

Huancayo (the same day)

	Mc/s	$>+1$	$>+\frac{1}{2}<+1$	$<-\frac{1}{2}>-1$	<-1	$>+\frac{1}{2}$	$<-\frac{1}{2}$
Washington	$>+1$	10	12	10	4	22	14
	<-1	25	33	13	8	58	21

Huancayo (day late)

	Mc/s	$>+1$	$>+\frac{1}{2}<+1$	$<-\frac{1}{2}>-1$	<-1	$>+\frac{1}{2}$	$<-\frac{1}{2}$
Washington	$>+1$	16	11	9	2	27	11
	<-1	17	33	19	7	50	26

EXPECTED

Huancayo

	Mc/s	$>+1$	<-1
Washington	$>+1$	9	7
	<-1	14	12

We note that days with depressed critical frequency occur at Huancayo much less frequently than would be expected by chance, and that Huancayo shows a positive deviation when the Washington deviation is negative.

Finally it is interesting to compare the deviations at two stations situated in opposite hemispheres but on the same longitude. For this purpose we have compared (Table 6) the results of observations made at Slough (51·5° N) and Johannesburg (26° S) during 1948 and 1949.

It will be observed that the degree of association is very small. There is a slight tendency for the simultaneous deviations of f_0F_2 at both stations to be abnormally infrequent and for the Johannesburg deviation to be positive when that at Slough is negative. This effect appears to be due to the influence of a few very large storms

Table 6. Comparison of F_2 Layer noon deviations at Johannesburg (26° S) and Slough (51° N) 1948 and 1949. (Same longitude, days of large deviations.)

OBSERVED

Slough

Johannesburg (same day)		+	−
	+	9	21
	−	21	12

Slough

Johannesburg (day late)		+	−
	+	9	23
	−	18	16

EXPECTED

Slough

Johannesburg		+	−
	+	14	14
	−	24	24

Slough

Johannesburg		+	−
	+	14	14
	−	24	24

(Days with deviation greater than 1 Mc/s above normal counted +; greater than 1 Mc/s below normal −.)

at Slough, the correlation being zero for the smaller disturbances. This effect is confirmed by the superposed-epoch analysis using days of negative deviation at Slough as zero days (see Fig. 4). This indicates a small delayed phase of positive deviations at Johannesburg lasting for about two days. A superposed-epoch curve for Johannesburg data, using days of negative deviations at that station as zero days, exhibits the usual positive phase followed by a longer negative phase. Thus this station behaves differently for days on which there is an association with a North Hemisphere storm and for those on which there is not. A comparison with the results from Graz (47° N, 15° E), which are less complete than the other data, suggests that a corresponding relation holds between these two stations *i.e.* when Johannesburg is showing a negative-type disturbance Graz often shows a positive or delayed positive effect. This suggests that the common positive-negative or negative type disturbances found at temperate latitudes are often associated with a purely positive disturbance at low latitude in the opposite hemisphere.

(c) *Variations of f_0F_2 with longitude*

An examination of all deviations for pairs of stations widely separated in longitude shows small and statistically insignificant associations. It is therefore necessary to select days for analysis on which there occur large deviations at one or more stations. A preliminary examination of such data shows that the number of days with significant negative deviations at a given station varies considerably with the longitude of that station. The total number of days with positive or negative deviations greater than a given minimum is found to vary slowly with longitude; though it increases considerably along a line of constant longitude from the equator towards the auroral zone.

As would be expected, the negative deviations of f_0F_2 which are associated with magnetic phenomena, appear to depend more on magnetic than on geographic latitude near the auroral zone. The curves of constant number of negative deviations

are not, however strictly parallel to magnetic latitude; *e.g.* Washington (Lat. 39° N; Long. 77° W) experiences more frequent and larger negative deviations than does Burghead (58° N, 3° W) or Leningrad (60° N, 30° E), although all three stations lie on the same magnetic latitude.

Some examples based on data for the year 1948, and illustrating these statements are given in Table 7.

Table 7.

No. of deviations exceeding ±1 Mc/s at several stations

Station	*Latitude*	*Longitude*	*Magnetic latitude*	*Negative deviations*	*All deviations*
Slough	51·5° N	0·°W	49° N	118	213
Mean of Ottawa and Churchill	52° N	85°W	71° N	155	239
Ottawa	45° N	76° W	63° N	144	211
Christchurch	43° S	173° E	50° S	124	218
Washington	39° N	77° W	55° N	126	192

VOL. 2 1952

We now examine the association between positive and negative deviations of f_0F_2 for stations at similar latitudes. If we consider the days on which a pair of stations both show deviations greater than 1 Mc/s, we find, even for widely separated stations, that for sites in the same hemisphere, there is a considerably greater degree of association than would be expected by chance. When, however, we analyse this association further we find that both its type and degree vary from pair to pair of stations.

Thus, for simultaneous large deviations at Ottawa and Slough, we have, Table 8 which does not indicate any appreciable association. If, however, we restrict the comparison to days of very great disturbance both at Ottawa and Slough we find, instead, the results set out in Table 9, from which it is clear that, for large deviations at Ottawa, the Slough deviation is more often positive than negative.

The same procedure restricted to days when Slough is very greatly disturbed shows that there is no reciprocal effect at Ottawa. This lack of reciprocity may well

Table 8. Comparison of F_2 Layer noon deviations at Ottawa and Slough 1948 and 1949. (Large deviations at both stations.)

OBSERVED

Ottawa \ Slough	+	−	
+	17	13	Total 81
−	24	27	

EXPECTED

Ottawa \ Slough	+	−	
+	15	17	Total 89
−	27	30	

(Deviations greater than 1 Mc/s above normal counted +; greater than 1 Mc/s below normal counted −.)

Table 9. Comparison of F_2 Layer noon deviations at Ottawa and Slough 1948 and 1949. (Large deviations at Slough, very large deviations at Ottawa.)

OBSERVED

Ottawa \ Slough	+	−	
+	8	5	Total 45
−	19	13	

EXPECTED

Ottawa \ Slough	+	−	
+	2	3	Total 18
−	6	7	

(Deviations greater than 1 Mc/s above normal counted + ; greater than 1 Mc/s below normal — for Slough.)
(Deviations greater than 1·5 Mc/s above normal counted + ; greater than 2 Mc/s below normal — for Ottawa.)

be an expression of the greater violence of the disturbances at Ottawa. The corresponding analysis for stations at temperate latitudes, north and south of the equator, does not indicate an association which is significantly greater than would be expected by chance. The results show that prolonged periods of great disturbance appear to involve both hemispheres; but, apart from the effects at low latitudes noted earlier, events in the two hemispehres do not appear to be directly connected.

In view of these large differences it is interesting to examine the degree of association between stations which are situated comparatively close together. As was the case for stations which were not widely separated in latitude, we find a large correlation between the deviations at pairs of stations, since in the majority of days similar phenomena are experienced at both stations. The deviations are, however, seldom identical at both stations, in either number or magnitude. This can be shown quite clearly with stations separated by 500 to 1000 km *e.g.* Lindau (Harz) (52° N, 10° E) is appreciably less affected by northern hemisphere storms than is Slough (51° N, 0° W); and Chungking (Lat. 29·5° N; Long. 106·5° E) has nearly twice as many large deviations as Delhi (Lat. 28·6° N; Long. 77° E).

It is interesting to note that McNish and Gautier [5], studying *all* f_0F_2 fluctuations at a large number of stations, found that the correlation coefficients for deviations decreased with increase in station separation.

(d) *The diurnal control of ionospheric storminess*

In order to demonstrate the characteristic trend of ionospheric storms throughout the hours of the day at a temperate latitude station the superposed-epoch method of displaying f_0F_2 data has been used. Days of marked depressions of noon f_0F_2 at Slough, associated with geomagnetic storm conditions, were selected as zero days and the hourly values of f_0F_2, expressed as fractions of the appropriate monthly median hourly values, extracted for each hour of zero day and of the days preceding and succeeding. The average results for 23 storms in the year 1949 are shown in Fig. 5, where it is seen that, subsequent to the small positive phase, the maximum of the negative phase is reached slightly before noon on the zero storm day, the value of the relative critical frequency recovering, thereafter, slowly to unity.

The same effect can be shown, in a different statistical fashion, by finding the ratios of the hourly values of f_0F_2 on magnetically disturbed days (D days) and on

magnetically quiet days (Q days). In Fig. 6 are shown the results of such an investigation for all months of the years of 1945 and 1949 at Washington. Here the ratio $f_0F_2(D)/f_0F_2(Q)$ is plotted for each hour of the local day, and, again, it is noted that the maximum relative departure of f_0F_2 from its quiet day value is reached just

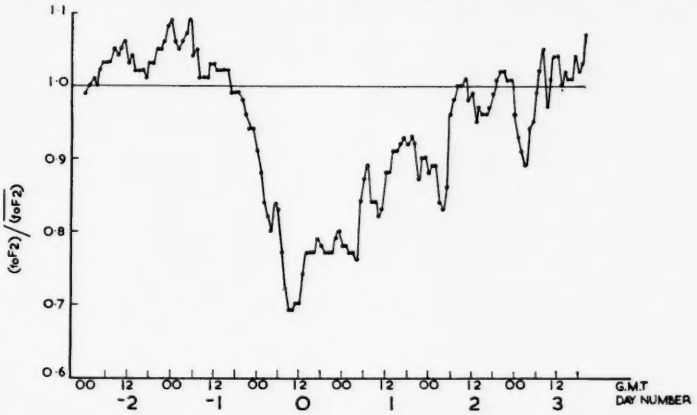


Fig. 5. Superposed epoch pattern of relative variation of f_0F_2 from its median value (23 Storms 1949).

VOL. 2 1952

before noon on disturbed days. A more detailed study shows that the general shape of the curve is not much affected by seasonal changes though the absolute magnitude of the relative depression of f_0F_2 is greatest in summer.

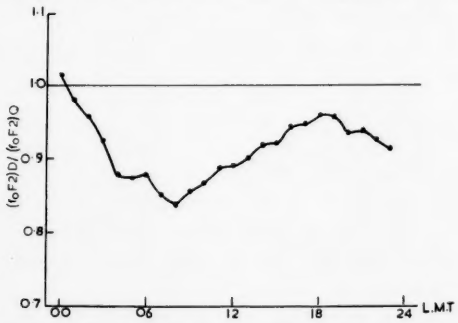


Fig. 6. Showing hourly values of the ratio of F_2 Layer critical frequencies on magnetically disturbed and quiet days at Washington, 1945 and 1949.

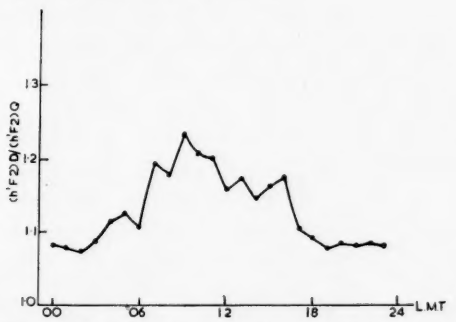


Fig. 7. Showing hourly values of the ratio of F_2 Layer equivalent heights on magnetically disturbed and quiet days at Washington, 1945 and 1949.

It is interesting to compare Fig. 6 with Fig. 7 in which the results of a corresponding comparison of $h'F_2$ (F_2 Layer equivalent height) on disturbed and quiet days is shown. The two curves are evidently closely related, the maximum relative value of $h'F_2$ being associated with the minimum relative value of f_0F_2. Also, a more detailed study shows, that the maximum relative values of $h'F_2$ are encountered in local summer.

Since phenomena such as have just been described indicate the pronounced diurnal control of ionospheric storminess and of the importance of local time in

correlating events at different stations, we have compared the ionospheric perturbations at three widely separated stations on the globe, Watheroo, Slough and Washington. The latitudes of all three are such that they all frequently experience negative phases during ionospheric storms. We have therefore made an examination of the times of the day at which negative phases are likely to start at the three stations. Results which are comparable in terms of local time have been found, in that negative phases start most frequently at all three stations when the local time is 0600 to 0700 h. There is also, at all stations, in the curve relating frequency of commencement of negative phase with local time, a less well-defined subsidiary maximum between 1800 and 2300 h. It is interesting to note, in the same connection, that negative phases have never been found to start between 1000 h to 1400 h local time.

(e) *The seasonal control of ionospheric storminess*

The fact that there is a seasonal control of ionospheric storminess at temperate latitude stations was shown by Appleton, Naismith and Ingram [2], and later

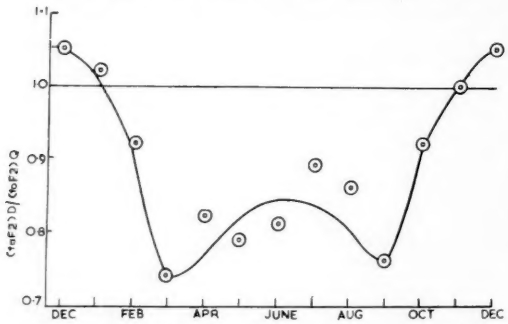


Fig. 8. Showing monthly values of the ratio of F_2 Layer critical frequencies on magnetically disturbed and quiet days at Washington 1944–1950.

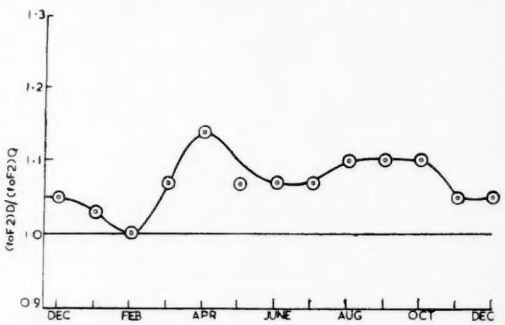


Fig. 9. Showing monthly values of the ratio of F_2 Layer critical frequencies on magnetically disturbed and quiet days at Huancayo 1938–1946.

confirmed by Berkner and Seaton [4]. The most marked depressions of f_0F_2, under ionospheric storm conditions, are encountered there in local summer, whereas, in winter, such departures from normality tend to be small, and more often than not positive in sense. Such effects can be strikingly illustrated by exhibiting the seasonal variation of the quantities $f_0F_2(D)/f_0F_2(Q)$ for noon conditions. The monthly values of these quantities for Washington, averaged over the years 1944–1950, are plotted in Fig. 8. This curve shows two interesting features. First, the fact that the major depressions on disturbed days occur in local summer, while, in December, the value of f_0F_2 at noon is slightly higher on disturbed days than on quiet days. Second, the well known equinoctial maxima of magnetic activity are paralleled in marked depressions of f_0F_2 on disturbed days at the same period of the year.

It was considered of interest to plot a similar curve for Huancayo (Fig. 9). Here it will be seen that the monthly values of noon $f_0F_2(D)/f_0F_2(Q)$ are all greater than unity.

But, perhaps, the most illuminating way of demonstrating the seasonal control of ionospheric storm manifestations is by reference to Fig. 10, which is a world chart showing the periods of marked depressions in f_0F_2 at noon during 1948. The disturbances are marked with lines the lengths of which are proportional to their severity. Here the following points are to be noted:

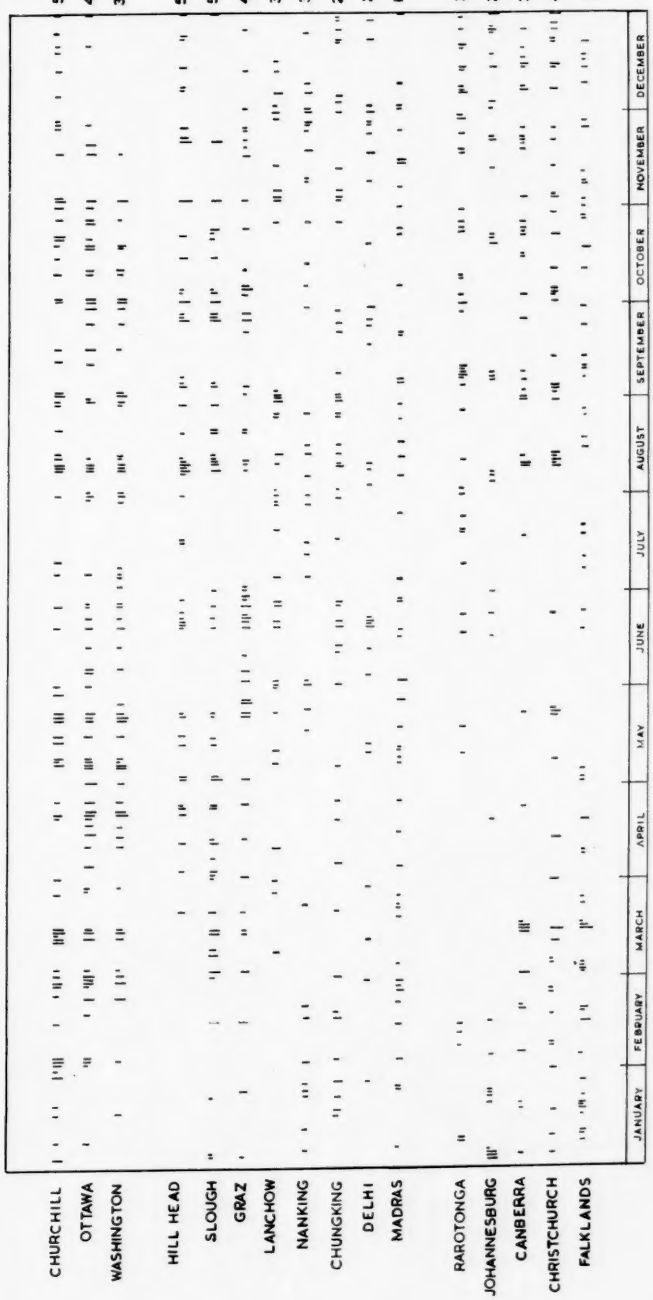


Fig. 10. F_2 disturbances at different stations 1948.

VOL. 2 1952

(a) Disturbances are relatively rare in local winter for the higher latitude stations, *e.g.* in December on the northern hemisphere and in June in the southern hemisphere.

(b) The same disturbances are evident on lines of constant longitude in a given hemisphere, at any rate for the higher latitudes.

(c) The major disturbances in one hemisphere are not evident as major disturbances in the other, suggesting that storms are not simultaneously developed around the two auroral zones.

(d) There is a decrease in the number and severity of disturbances experienced as one proceeds from high to low latitudes.

4. Further Discussion of the Statistical Results

The investigations described above show that, although certain characteristic modes of perturbation may be readily recognised during periods of ionospheric disturbance, the general morphology of the phenomena is complicated. The results of the superposed-epoch analysis at suitable pairs of stations, *e.g.* Slough and Johannesburg, show also that it is possible for at least two alternative modes of perturbation to be present at a given station. It is, therefore, particularly important to distinguish between statistical results which indicate the absence of correlated perturbations at different stations and those which indicate the association of different types of response.

Considering the results from a pair of stations in the same hemisphere which show similar total disturbance activity and modes of perturbation, *e.g.* Slough and Chungking, we find that the number of coincidences within plus or minus one day of large negative fluctuations at noon is only slightly greater than would be given by chance. Typical analysis indicates that at least 85 % of these disturbances can be regarded, from statistical evidence alone, as being limited to a restricted range of longitudes.

When the results for pairs of stations in opposite hemispheres are compared the apparent correlation will, of course, be reduced by the seasonal variations of the type and magnitude of the perturbations. Even in sunspot maximum years, when the duration of the positive phase seldom exceeds one day, and hence the presence of a biphase disturbance does not affect the apparent correlation between large negative deviations, the correlation found is insignificant. We may conclude, then, that it is probable that the great majority of disturbances are essentially localised and unipolar.

It is interesting to examine the implications of this conclusion by reference to the relation between ionospheric and magnetic disturbances. For this purpose we have examined the ionospheric phenomena at several stations during and immediately after the beginning of isolated periods of magnetic activity. Unfortunately, isolated magnetic perturbations are only common in the years of minimum solar activity during which the biphase type of perturbance in temperate latitudes is particularly frequent and most highly developed. Thus the results are not entirely representative of all years. We have used data from Slough, Burghead (Scotland) and Huancayo, for 85 isolated periods of magnetic activity at Abinger during 1942 and 1943, and denoted a negative deviation by N, a positive by P, no effect by O, a typical biphase perturbation by P/N and delayed positive or negative changes by O/P and O/N

respectively. We then find that the numbers of times on which the fluctuations of the noon f_0F_2 parameter were of a given type are as shown in Table 10.

It should be noted that, as there is always some day-to-day variation in the f_0F_2 parameter and the minimum significant deviation, $\pm 0{\cdot}5$ Mc/s, is small, this type of analysis will tend to underestimate the number of occasions on which no significant change was present.

The increase in the number of negative deviations and the decrease in the number of biphase disturbances with the increase of latitude as we go from Slough (Lat. $51\frac{1}{2}°$) to Burghead (Lat. 56°) are very characteristic, and similar results are obtained using data from other comparable pairs of stations. It will be noted that the Huancayo figures confirm our earlier conclusions concerning the equatorial positive mode of perturbation. Moreover the analysis disclosed that remarkably few of the periods of negligible variations were coincident at Huancayo and Slough. The comparatively large number of periods showing an equatorial positive phase associated with periods of magnetic activity is consistent with the results of other statistical investigations. The whole evidence therefore suggests that the equatorial type of positive perturbation is found over a much wider range of longitudes than is usual for the other typical types of perturbation.

Table 10.

	P/N	*O/N*	*N*	*P*	*O/P*	*O*	*Total*
Burghead .	26	9	23	9	—	13	80
Slough. . .	35	9	16	9	—	16	85
Huancayo .	—	—	—	30	29	24	83

(Note that at Burghead there were 5 occasions, and at Huancayo 2 occasions, when no ionospheric data were available.)

VOL. 2 1952

It is interesting, finally, to consider the F_2 Layer fluctuations at latitudes higher than the auroral zone. The data available for this area are rather meagre since screening effects due to storm perturbations in the E and D Layers are very common.

The main feature of the perturbations is the great violence and irregularity of the changes in layer height and configuration during periods of magnetic disturbance. These differ, however, from the rather similar perturbations in the auroral zone itself in that positive and negative fluctuations of f_0F_2 are almost equally large and frequent. Thus the distribution curve of the fluctuation of f_0F_2 is symmetrical about the median or mean value near the magnetic pole, whereas in the auroral zone and temperate latitudes, it is very skew, negative changes relative to the median being larger than the positive fluctuations.

Summarising the main results of our statistical survey, we can generalise our conception of ionospheric storm perturbations at different seasons and then compare them, as we shall in Part II of this paper, with events during a few selected individual storms.

A typical ionospheric storm perturbation appears to be localised and substantially unipolar, though it may extend beyond the equator. It appears to be initiated by a magnetic perturbation but to be controlled mainly by terrestrial geographic and geomagnetic factors.

The sequences of events at stations situated along a geographic or, probably more accurately, magnetic meridian, are often closely connected and show characteristic changes which vary with latitude and, in general, with season. At high latitudes, in the auroral zone, there is a close connection between magnetic perturbations and disturbances in the F_2 region, the critical frequency being greatly

decreased and the height of the layer greatly increased during the periods of magnetic activity. At lower latitudes, however, these fluctuations are replaced by more regular and longer perturbations which often show a characteristic bi-phase structure. This structure is most evident in winter and in the years of minimum solar activity. It consists of a positive phase, during which the critical frequency is above normal, followed by a negative phase in which the critical frequency is suddenly greatly depressed and takes between one and several days to return to normal.

The change-over between positive and negative phases can only occur between certain hours L.M.T. the most common time being about 0600 L.M.T. for small storms and either near 0600 L.M.T. or 1800 L.M.T. for large storms. In summer months the positive phase is greatly shortened or absent altogether but the regular negative phase is more pronounced and of longer duration.

At still lower latitudes the positive phase increases in duration and the negative phase decreases in magnitude until, eventually, only a positive phase is left. This positive phase is often most marked one or two days later than the associated phenomena at high latitudes.

Positive phase phenomena, in particular the occurrence of high noon values of f_0F_2, are also occasionally present at high and temperate latitudes during severe disturbances outside but probably chiefly near the edge of the main zones of negative and bi-phasal disturbance.

We thus form a picture of a complex perturbation with magnetically controlled fluctuations replaced by a bi-phase and finally positive phase perturbation as we move to lower latitudes. The position and size of this perturbation varies from storm to storm so that phenomena characteristic of widely separated parts of a typical perturbation may be observed on different occasions at a given station. There are also characteristic seasonal and diurnal variations in the magnitude and type of disturbance.

After allowing for these effects we find that the number and intensity of the negative and bi-phase disturbances vary considerably with longitude, but that there is no evidence of any significant difference in activity between the North and South hemispheres.

Acknowledgements—We wish to acknowledge gratefully the assistance we have received from the many organisations and individuals who have furnished us with ionospheric data. We also wish to thank Mrs. A. G. Pritchard, of *Edinburgh University*, and Mrs. I. Moorat, of the *Radio Research Station*, Slough, for assistance in the preparation of the paper. The work described above was carried out as part of the programme of the *Radio Research Board*. This paper is published by permission of the Director of Radio Research of the *Department of Scientific and Industrial Research*.

References

[1] Appleton, E. V. and Ingram, L. J.; Magnetic Storms and Upper-Atmosphere Ionisation. Nature, Lond. 1935 **136** 548–549. [2] Appleton, E. V., Naismith, R. and Ingram, L. J.; British Radio Observations During the Second International Polar Year 1932–1933. Phil. Trans. Roy. Soc. A 1936 **236** 191–259. [3] Harang, L.; Änderungen der Ionisation der höchsten Atmosphärenschichten während der Nordlichter und erdmagnetischen Störungen. Geofysiske Publikasjoner 1937 **XI** (No. 17) Utgitt Av Det Norske Videnskaps-Akademi I Oslo. [4] Berkner, L. V. and Seaton, S. L.; Systematic Ionospheric Changes Associated with Geomagnetic Activity. Terr. Mag. 1940 **45** 419–423. [5] McNish, A. G. and Gautier, T.; Ionospheric Radio Propagation. Nat. Bureau Stand. Circ. 462, 1948, p. 61.

RESEARCH NOTES

On negative ions of molecular oxygen in the *D*-layer

In a paper [1] published in this journal a short time ago the authors discussed the possible concentration of O_2^- in the *D*-layer. The values adopted for the attachment coefficient, η, were based on the experimental work of BRADBURY and of HEALY and KIRKPATRICK [2]. They were, however, regarded as upper limits since theoretical investigation (BATES and MASSEY [3]), failed to reveal any mechanism which could be as rapid as implied by them. Support for this attitude is provided by BIONDI's recent measurements on the rate of attachment of slow electrons to molecular oxygen [4]. According to these η is only some $1{\cdot}4 \times 10^{-15}$ cm^3/sec through a pressure range of from 8 to 25 mm of mercury. This is several orders of magnitude smaller than the values found earlier and, as BIONDI points out, might well be accounted for by the simple radiative process

$$O_2 + e \rightarrow O_2^- + h\nu$$

(though other processes cannot be excluded). The cause of the difference is obscure and further laboratory studies are required. If these confirm BIONDI's results the original estimates of λ, the day negative ion to electron ratio, would require to be revised downwards.

The following table gives the calculated values of λ at the 75 and 90 km levels taking η to be $1{\cdot}4 \times 10^{-15}$ cm^3/sec, and the destruction of the O_2^- to be due mainly to photo-detachment. As the rate coefficient of this latter process is still uncertain two alternative assumptions were made: (A) that it is 5×10^{-1}/sec and (C) that it is 5×10^{-3}/sec. The basis of these was discussed in the previous paper [1]. We need only mention here that (A) is considered as likely to be of the correct order of magnitude; and that (C) is an extreme figure which general quantal considerations indicate is only barely credible, and which indeed is certainly much too low if the process observed by BIONDI is the inverse of photo-detachment (that is if it is radiative attachment to the ground state).

Altitude	*Calculated* λ	
75 km	(A)	0·5
	(C)	50
90 km	(A)	0·05
	(C)	5·0

As can be seen the new upper limits to λ [provided by case (C)] are far less than those given earlier (*cf.* Table 3 of [1]). They are actually such that COWLING's formula [5] gives the contribution of the negative ions to the transverse conductivity of *D* to be *much smaller than the contribution of the free electrons.* Thus BIONDI's determination of η increases the already large discrepancy between the transverse conductivity which is predicted and that which is deduced from the diurnal magnetic variations; but to some extent it eases the task of the theorists, for, if correct, it renders the uncertainty in λ relatively unimportant. An independent check on it is urgently needed.

Attention may also be drawn to the fact that if the value of λ in *D* should indeed prove smaller than has hitherto been supposed the estimates of the ionization rate would have to be correspondingly reduced. Moreover, the assumption (*cf.* MASSEY, BATES and SEATON [6]) that equilibrium in the layer is preserved by ionic recombination would require re-examination. In case (A), for example, the effective recombination coefficient at the 75 and 90 km levels would only be some 5×10^{-8} and 5×10^{-9} cm^3/sec respectively, even if the ionic recombination coefficient were as high as 10^{-7} cm^3/sec. Consideration must therefore be given to the possibility that dissociative recombination is the process operative.

References

[1] BATES, D. R. and MASSEY, H. S. W.; J. Atmosph. and Terr. Phys. 1951 **2** 1. [2] BRADBURY, N. E.; Phys. Rev. 1933 **44** 883. HEALEY, R. H. and READ, J. W.; Behaviour of Slow Electrons in Gases,

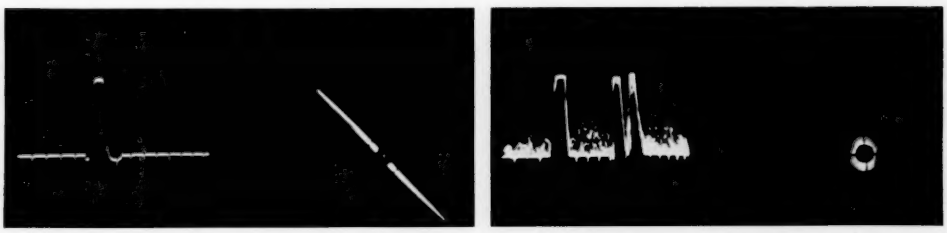
VOL. 2 1952

Amalgamated Wireless Ltd., Australasia, 1941. [3] BATES, D. R. and MASSEY, H. S. W.; Phil. Trans. Roy. Soc. A 1943 **239** 269. [4] BIONDI, M. A.; Phys. Rev. 1951 **84** 1072. [5] COWLING, T. G.; Proc. Roy. Soc. A 1945 **183** 453. [6] MASSEY, H. S. W.; Negative Ions (2nd edition). Cambridge University Press, 1950. BATES, D. R. and SEATON, M. S.; Proc. Phys. Soc. B 1950 **63** 129.

Queen's University, Belfast. D. R. BATES.
University College, London. H. S. W. MASSEY, F.R.S.

Polarization of radio waves reflected from the ionosphere

Measurements of the limiting polarization of radio waves reflected from the ionosphere at vertical incidence were made at the Norwegian Defence Research Establishment at Kjeller near Oslo (60° N, 11° E) during November 1950 and at the Auroral Observatory in Tromsø (70° N, 19° E) during September and October 1951.



a *b* *a* *b*

Fig. 1. Left: *a* shows the ground-ray with height marks each 50 km; *b* the ground-ray is gated and displayed on a second CRT.
Right: *a* shows F_2-echoes; *b* the x-component is gated and displayed on a second CRT.

The experimental system for the measurements was the standard system using crossed loops, separate receiver channels with a common oscillator and an oscilloscopic representation of the ellipse of polarization.

The following results were obtained:

1. The sense of rotation for the third magneto-ionic component (the z-component) was determined. The z-component was found to be polarized in the same way as the ordinary component. This result has also recently been obtained by HOGARTH (Nature, Lond. 1951 **161** 597).

2. At Kjeller, absorption measurements were carried out by F. LIED at the same time as the polarization measurements were made. Correlation was found between the mean values of the absorption index D and the axis ratio ε of the ellipse of polarization. The correlation is shown in Fig. 2. The observations were made at about 12 h M.E.T. every day on a frequency of 3·6 Mc/s. On this frequency only the ordinary magneto-ionic component was present.

This result seems to be in discrepancy with the magneto-ionic theory. According to the ray theory an increase in the absorption should make the ellipse of polarization more circular than eccentric. The effect may be explained as due to a resplitting of the reflected ordinary wave into both an ordinary and an extraordinary wave during the passage through the lower part of the ionosphere.

3. During polar magnetic storms and auroral displays in Tromsø the ellipse of polarization was found to undergo marked variations. In Fig. 3 mean values of the axis ratio ε of the ellipse of polarization are plotted together with a record of the horizontal earth magnetic force on a magnetically *quiet* day and a magnetically *disturbed* day. On 23rd October 1951, at about 16 h M.E.T., a polar magnetic storm appeared. At the same time, in less than one hour, the axis ratio ε changed from 0·92 to 0·68.

Further studies are required before the observed effect can be fully explained, but the effect may well be due to a mechanism similar to the one proposed in (2). It is well known that, during polar magnetic storms, the concentration of electrons in the lower parts of the ionosphere

is increased. During such conditions the occurrence of discontinuities, and thereby the conditions for a resplitting of the ordinary wave into an ordinary and an extraordinary wave, may be expected.

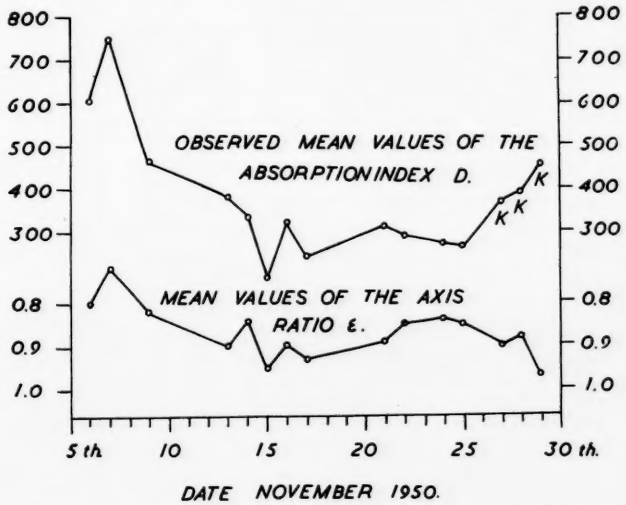


Fig. 2. Observed mean values of the absorption index D and the axis ratio of the ellipse of polarization. K is a symbol for storm conditions. The observations were made at Kjeller.

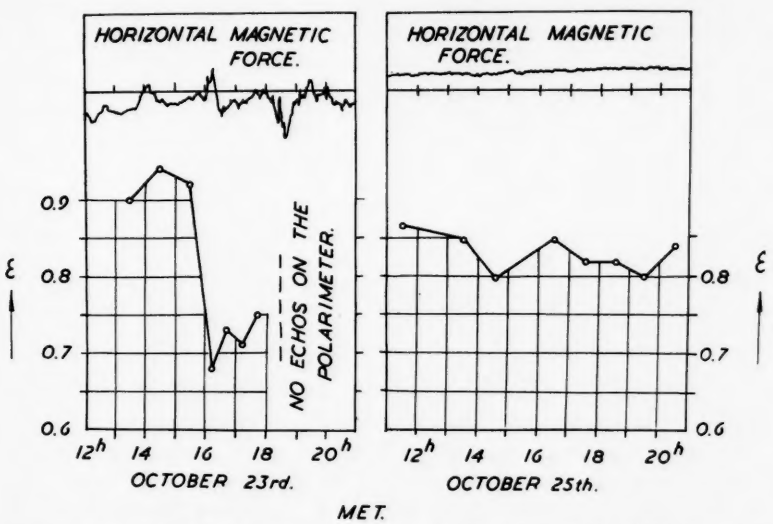


Fig. 3. Observed mean values of the axis ratio ε of the ellipse of polarization for an earth-magnetically quiet and a disturbed day. The observations were made at Tromsø.

The author would like to express his indebtedness to Dr. L. Harang, under whose supervision the present work was done. He also would like to thank Mr. F. Lied for his most inspiring help.

Norwegian Defence Research Establishment, Kjeller, Oslo. Bjørn Landmark.

VOL. 2 1952

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 256 to 258. Pergamon Press Ltd., London

Indices of geomagnetic activity

of the Observatories ABINGER (*Ab*), ESKDALEMUIR (*Es*) and LERWICK (*Le*)

January to March 1952

The figures given on pages 256 to 258 represent the *K*-indices for three-hour intervals, beginning with 00–03 hrs for the first and ending with 21–24 hrs for the eighth figure.

January 1952

Day	*Ab* *Range for K = 9: 500γ*			*Es* *Range for K = 9: 750γ*			*Le* *Range for K = 9: 1000γ*			*Day*
	K-Indices		*Sum*	*K-Indices*		*Sum*	*K-Indices*		*Sum*	
1	4233	3332	23	4233	3232	22	4223	2232	20	1
2	2222	2232	17	2212	1232	15	2211	1232	14	2
3	2122	3143	18	1111	1033	11	1111	1033	11	3
4	3113	3335	22	3212	3334	21	3212	2235	20	4
5	3454	4435	32	3444	4335	30	3543	4435	31	5
6	4444	3532	29	4433	2432	25	4443	2442	27	6
7	2233	4444	26	2222	3444	23	2222	3444	23	7
8	3323	2321	19	3322	2320	17	3322	2321	18	8
9	0332	2343	20	0221	2242	15	0221	1242	14	9
10	3433	3535	29	2433	2534	26	3432	2435	26	10
11	3333	4544	29	3322	3544	26	4322	3543	26	11
12	5433	3555	33	4433	3454	30	4433	3455	31	12
13	3334	5455	32	3334	4445	30	3323	5455	30	13
14	4344	4545	33	4333	4544	30	4433	3455	31	14
15	3433	6443	30	3333	5433	27	3333	5543	29	15
16	3123	2343	21	3122	2333	19	3122	2233	18	16
17	2122	1132	14	1221	1122	12	2121	1132	13	17
18	1112	1111	9	1100	0101	4	2111	1111	9	18
19	1121	2232	14	0100	1222	8	1011	1132	10	19
20	0111	0133	10	0211	1132	11	1111	1133	12	20
21	2222	1122	14	2101	1022	9	2212	1122	13	21
22	2121	1233	15	2111	1123	12	2112	1233	15	22
23	3122	4454	25	3122	3444	23	3112	3445	23	23
24	3123	3433	22	3211	2323	17	3111	2333	17	24
25	3212	3333	20	3112	2332	17	3212	2332	18	25
26	0022	3210	10	0012	2110	7	0012	2211	9	26
27	1333	5544	28	0333	4544	26	1333	5654	30	27
28	3433	4424	27	2433	3323	23	3332	3423	23	28
29	2123	3566	28	2022	3566	26	2022	4685	29	29
30	3333	3133	22	3322	2133	19	3332	2133	20	30
31	1122	3323	17	1111	3313	14	1111	3433	17	31

February 1952

Day	*Ab* Range for $K = 9$: 500γ		Sum	*Es* Range for $K = 9$: 750γ		Sum	*Le* Range for $K = 9$: 1000γ		Sum	Day
	K-Indices			K-Indices			K-Indices			
1	3244	4455	31	3134	4454	28	3233	3455	28	1
2	2332	2323	20	3322	2222	18	3321	1222	16	2
3	2112	1121	11	1110	2121	9	2111	1121	10	3
4	2111	2110	9	2111	2110	9	1111	1111	8	4
5	0111	1101	6	0001	1000	2	0111	1001	5	5
6	1234	4566	31	1223	3666	29	1223	3767	31	6
7	4345	4455	34	3344	4354	30	5433	3345	30	7
8	4354	4555	35	4343	4454	31	5343	3444	30	8
9	3323	4454	28	3323	3454	27	4422	3455	29	9
10	3223	2466	28	2222	2456	25	3222	2457	27	10
11	5434	3355	32	5324	3355	30	6423	2355	30	11
12	5433	4455	33	4433	4345	30	5433	4355	32	12
13	4333	4446	31	4232	4445	28	4222	4445	27	13
14	3233	3343	24	3112	2343	19	3222	2443	22	14
15	2233	2213	18	1222	1203	13	2222	1203	14	15
16	3454	5455	35	3454	4455	34	3454	5454	34	16
17	4223	3222	20	4212	2211	15	6312	2212	19	17
18	3223	2244	22	3213	2133	18	4213	2234	21	18
19	5343	3465	33	4333	3355	29	5532	3365	32	19
20	3333	3311	20	2232	1311	15	3232	1312	17	20
21	1122	3222	15	1110	2212	10	1111	2122	11	21
22	0132	3232	16	1021	2132	12	1010	2132	10	22
23	2133	2314	19	2110	0214	11	2111	1104	11	23
24	6635	5564	40	6635	5564	40	7734	4565	41	24
25	4133	5231	22	4123	4121	18	5323	3121	20	25
26	4334	3445	30	3233	2445	26	3222	2545	25	26
27	4443	2455	31	4343	3455	31	4443	3455	32	27
28	4345	4452	31	4344	3452	29	4443	3452	29	28
29	2333	3424	24	2333	4323	23	2222	3323	19	29

March 1952

Day	K-Indices		Sum	K-Indices		Sum	K-Indices		Sum	Day
1	4332	4311	21	3322	3311	18	3322	3311	18	1
2	1212	2111	11	1201	2100	7	1211	1110	8	2
3	1032	5565	27	0022	4465	23	0121	4578	28	3
4	5545	4555	38	4544	4555	36	6645	4566	42	4
5	5545	5656	41	5545	4557	40	7554	6769	49	5
6	7654	4355	39	8653	4345	38	9763	3355	41	6
7	3435	5656	37	3435	5546	35	3434	5556	35	7
8	5333	4564	33	5333	4564	33	5433	4556	35	8
9	5334	4555	34	5334	4546	34	6434	3558	38	9
10	5343	4655	35	4333	4555	32	5333	4557	35	10
11	4333	4564	32	4332	3453	27	7432	4354	32	11
12	4233	3443	26	3323	3443	25	5422	4343	27	12
13	3133	4224	22	2122	3224	18	3122	2224	18	13
14	2122	2112	13	2112	2102	11	2112	2212	13	14
15	2232	3344	23	2221	3344	21	2221	3344	21	15
16	1333	2344	23	1333	2334	22	2233	2334	22	16

VOL. 2 1952

March 1952 (Continued)

Day	*Ab* Range for $K = 9: 500\gamma$		Sum	*Es* Range for $K = 9: 750\gamma$		Sum	*Le* Range for $K = 9: 1000\gamma$		Sum	Day
	K-Indices			*K-Indices*			*K-Indices*			
17	3433	4433	27	3333	4433	26	4333	3433	26	17
18	2233	3413	21	1222	2303	15	2222	2312	16	18
19	2122	2333	18	2021	2232	14	2011	2233	14	19
20	0122	1121	10	1112	1111	9	1111	1111	8	20
21	3333	5445	30	2333	4435	27	3233	4436	28	21
22	4433	2355	29	3333	2355	27	5442	2347	31	22
23	6243	2464	31	7333	3454	32	7443	2353	31	23
24	4444	5444	33	4433	5444	31	3533	4454	31	24
25	2343	4344	27	2243	4333	24	3233	3333	23	25
26	4321	3343	23	4231	3242	21	4412	3243	23	26
27	2333	3433	24	2332	3423	22	2322	3323	20	27
28	2221	2221	14	2110	2010	7	2111	2111	10	28
29	1123	3234	19	0112	2223	13	0112	2223	13	29
30	3322	4566	31	3222	4456	28	3321	3558	30	30
31	6554	4654	39	7554	4554	39	7643	3556	39	31

CORRECTION

In the paper entitled "An apparatus for simultaneous registration of potential gradient and air-earth current" by H. W. KASEMIR, in this Journal, Vol. 2, p. 32-37, the caption to Fig. 4 should have read as follows:

Fig. 4. Curve I: Diurnal variation of the thunderstorm activity (F. J. W. WIPPLE and F. J. SCRASE), Curve II: Diurnal variation of the air-earth current on the Nebelhorn.

Dr. BERNARD LYOT†

The sudden death on 2 April of Dr. BERNARD LYOT at the early age of 55 years is a severe loss to astronomy. Dr. LYOT had gone to the Sudan to observe the total eclipse of the Sun on 25 February, near Khartoum. After the eclipse he was spending some weeks in Cairo studying at the Helwan Observatory the data that he had obtained; he died from heart failure while in the train on the way from Cairo to Helwan.

LYOT was born in Paris on 27 February 1897. He studied engineering at the Electrical High School. After receiving his diploma in 1917 he worked in the laboratory of General FERRIÉ on various war problems; there he attracted the attention of Professor PÉROT, who was associated as a physicist with the Meudon Observatory. LYOT had always had an interest in astronomy; at the age of 16 he had built a small Observatory at the top of the family mansion in Paris. PÉROT introduced LYOT to the Meudon Observatory and strongly recommended him to the Director, HENRI DESLANDRES, with the result that LYOT was appointed to the staff. The whole of his subsequent career was spent in the service of the Observatory, at Meudon or at its mountain stations.

The first big problem to engage LYOT's attention was the polarisation of the light from the Moon and planets. For this purpose he devised and constructed a polarimeter that was ten times more sensitive than any previously made, with which he was able to detect a polarisation as small as 0·1 per cent. He studied the amount of the polarisation and the direction of its plane for the various bodies at different phase angles, and showed that at a certain angle of diffusion, which he called the inversion angle, the polarisation passed through a zero value and its plane changed by 90°. At the same time he studied in the laboratory the polarisation of light diffused by a great variety of terrestrial substances, and by particles in suspension in gaseous media, in order to compare with those the results for the Moon and planets. In that way he was able to form some tentative conclusions about the nature of the surfaces of the Moon and Mercury, and about the atmospheres of the other planets.

VOL. 2 1952

But LYOT's greatest work was associated with the study of the Sun's corona. Some astronomers had proved to their own satisfaction that the observation of the corona without a total eclipse of the Sun was impossible. LYOT did not accept this view, and he proved its falsity by the ingenious design of the instrument now known as the coronagraph. With this instrument an artificial eclipse of the Sun is produced; by suitable design, scattered and diffracted light is prevented from reaching the field of view. Special care was taken in the optical work, so that light would not be diffused or scattered by scratches, striations, bubbles etc., on or in the glass. The coronagraph was installed on the Pic du Midi, at an altitude of 1400 metres, above the permanent dust layer, where atmospheric diffusion was small. The first photograph of the spectrum of the corona without an eclipse was obtained on 8 August 1930. In subsequent years, successive improvements to the telescope and spectrograph were made, and the corona was brought under fairly continuous observation.

In this work LYOT obtained photographs of the corona and of its spectrum. He made accurate wave-length measurements and discovered several hitherto unknown lines, including two strong lines in the infra-red. He proved that all the lines are about 1 A in width, showed that they were not all distributed in the same way round the Sun, and observed their relative variations in intensity. He also studied the prominences, their spectra, and their motions. He was the pioneer in obtaining records of the changes in prominences—their evolution and dissipation—by means of cinematograph films, which has revolutionised the study of the phenomena that they present.

For the study of differences in distribution of the various coronal radiations, photography of the corona in the light of individual radiations was desirable. To meet this need LYOT developed and constructed a new type of polarising filter, with a high percentage transmission for a narrow band of less than 1 A in width. With such a filter LYOT was able to photograph the Sun in practically monochromatic light with an exposure time of less than one second, with results superior to those obtained with the spectroheliograph. With a modified form of the filter, he obtained six practically monochromatic transmissions, coinciding closely with the red coronal line, 6375 A, the green coronal line, 5303 A, and four important chromospheric radiations (Hα, D_3 of helium, C of magnesium, and Hβ). He also obtained simultaneous cinematograph records of solar prominences in Hα, 6375 A, and 5303 A.

But LYOT has recently gone further. By a remarkable tour-de-force he succeeded in measuring at Meudon the intensity of the green coronal line, even through thin cirrus cloud, and without using carefully selected optical components. The method was to isolate a narrow spectral band containing the green coronal line and to compare it with an adjacent band of the same global intensity. By

means of a special photoelectric polarising photometer, the difference, attributable to the coronal radiation, could be measured. Comparison of the intensity contours of the line round the Sun's disk with coronagraph observations at the Pic du Midi confirmed the reality of the observed variations. LYOT had planned to attack some even more difficult problems for which, he said, he would need ten years. But what he had in mind is not known.

LYOT was elected an Associate of the Royal Astronomical Society in 1937 and was awarded its Gold Medal in 1939. In that year he was elected a member of the French Académie des Sciences. He was awarded the Bruce Gold Medal of the Astronomical Society of the Pacific in 1947. A man of a charming and modest disposition, he had many friends in this country.

H. S. J.

Book review

MOXON, L. A.; **Recent advances in radio receivers.** Pp. x + 183. Cambridge University Press, 1949. 18 s. net.

This book is one of a series of monographs describing the advances made in radio technique during the decade which began with the outbreak of war in 1939. While the field covered by the title is a wide one, the scope of the book has been deliberately limited by the author's interest in the development and application of radio receivers for use at frequencies above 100 Mc/sec. The principles described are, however, by no means restricted to very high frequency technique, as one chapter is devoted to intermediate frequency amplifiers, and this is followed by another on practical receiver design, in which some of the basic characteristics of broadcasting and communication receivers are discussed.

After a brief introduction, four chapters comprising nearly half the book, are devoted to a discussion of the concept of noise factor, as a criterion of assessing the performance of a receiver; to the principles of design of a receiver, including the mixer unit, to give a minimum value of this noise factor; and to a discussion of the methods by which this overall performance can be measured. These chapters are most important and instructive, since the whole problem of success in radio reception consists in producing at the output terminals of the equipment a signal which can be distinguished, aurally or visually, from the accompanying noise. Furthermore, as the author describes, the fact that noise sources are actually used as a means of measuring receiver performance illustrates the manner in which the radio worker has turned to advantage a phenomenon which was at first regarded merely as a cause of interference or annoyance. Although the principles discussed are of general application, much of the work concerned was originally done on radar receivers in which it was necessary to effect a compromise between a large band-width and a low noise factor. The uses of feed-back, the grounded-grid triode and the modern high-slope pentode, in radio frequency amplifiers are discussed, and their performance compared at various frequencies up to 600 Mc/sec. The chapter on intermediate frequency amplifiers is concerned with the developments in technique which have resulted from the demand for efficient amplification over band-widths of one or more megacycles per second for such purposes as radar and television. By the use of coupled, staggered amplifier circuits in cascade, a sufficient uniformity of gain and time delay can be achieved, so that pulse-modulated signals can be amplified without distortion. Two later chapters deal with the trends in design of new types of receiver for communication, television and radar purposes; while the ninth and last chapter describes a number of circuit devices used for various purposes in receiver design technique.

Altogether the author seems to have achieved his objective in the production of a most useful guide to the principles of design of radio receivers, and this monograph can be recommended particularly to those concerned with pulse reception of which radar and television form two notable applications.

R. L. SMITH-ROSE

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 261 to 265. Pergamon Press Ltd., London

The graphical representation of the longitude effect in F_2-region

C. M. Minnis, M.Sc., F. Inst. P.

Communication from the National Physical Laboratory

(Received 11 January 1952)

ABSTRACT

Assuming that the type of fF_2/dip angle relation suggested by Appleton is valid, a network of geographical zones can be constructed within each of which the change of fF_2 along a line of latitude does not exceed some upper limit. A typical set of zones has been constructed but the boundaries are too complex and variable to be used in practice for the accurate representation of the geographical distribution of fF_2.

It is concluded that for commercial use the best method would be to make use of a set of hourly charts each of which would show the distribution at a fixed U.T.

VOL.
2
1952

Introduction

It is well known that the critical frequency for either the E- or F_1-region (fE, fF_1) for a given U.T. can be expressed to a high degree of accuracy as a function of the solar zenith angle alone. It follows from this that a contour map showing the world distribution of fE or fF_1 at a given U.T. can also be used to determine the distribution at another U.T. by a simple shift of the longitude scale. In particular, fE and fF_1 along a given line of latitude are functions only of L.M.T. and not of longitude.

For F_2-region, however, fF_2 along a given line of latitude is not only a function of L.M.T. but of longitude also.

The Present Zone System for F_2-Region

The fact that fF_2 depends on both L.M.T. and longitude implies that the world distribution of fF_2 at all L.M.T.'s cannot be represented with sufficient accuracy on a single contour map. It has been known for some years that this "longitude effect" is connected with the fact that the geomagnetic and geographic equators are not parallel [1]. For this reason it is customary at present to divide up the earth's surface into three zones [2] within each of which the change of geomagnetic latitude for a given geographic latitude is not too great. It is then assumed that for all points within each zone the change in fF_2 with change of longitude is sufficiently small to permit the determination of fF_2 at any L.M.T. from a single set of contours in the same way as for E- and F_1-regions.

Doubt has been expressed in various quarters as to the validity of these assumptions for the present zone system and Oboril and Rawer [3] have suggested a revision of the zone boundaries. Such proposals are usually made with the intention of improving the accuracy of determination of fF_2 in some specific area, and it does not necessarily follow that the accuracy in other areas will not suffer. Consequently it seems desirable to make a fresh approach to the problem without any preconceived ideas about where the zone boundaries should lie. It will be assumed that the boundaries should be chosen so that each zone includes an area of the

earth's surface within which the range of values of fF_2 for any given latitude and L.M.T. does not exceed some arbitrarily chosen limit. Providing this condition can be satisfied, the way in which fF_2 changes along a line of longitude is immaterial.

Relation between fF_2 and Dip Angle

In recent years it has become clear that the change in fF_2 along a line of latitude at a fixed L.M.T. depends more closely on the magnetic dip angle than on the true geomagnetic latitude. This is fully discussed by Appleton [4] and by Lépéchinsky and Pillet [5] and for the purposes of this note it will be assumed that for a given epoch and L.M.T., the relation between fF_2 and dip angle

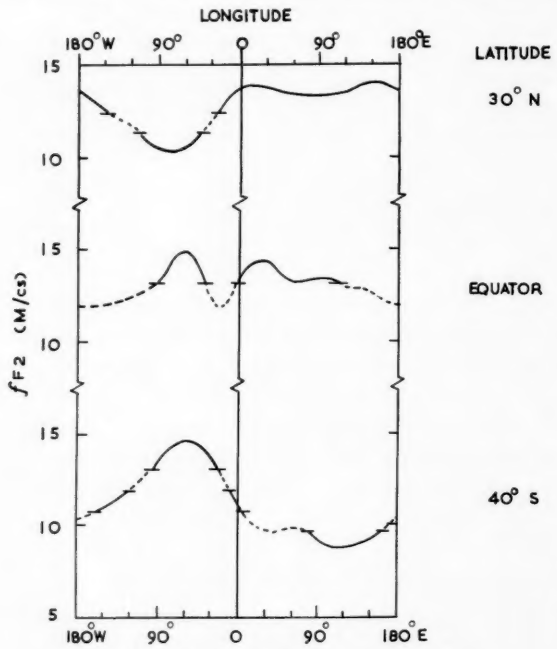


Fig. 1. Relation between fF_2 and longitude (noon, September 1948).

Table 1.
Noon, September 1948

fF_2 (Mc/sec)	Dip Angle (deg)
12·0	0
12·5	10
13·5	20
15·0	30
14·0	40
12·5	50
11·0	60
9·0	70
7·0	80
5·5	90

can be represented by a smooth curve. The curve used here for illustrative purposes is that for noon in September 1948 (extrapolated to 90°) and is represented by the values in Table 1. Given this relation and the geographical distribution of dip angle, it is possible to construct a family of curves showing the changes in fF_2 along lines of latitude: this has been done at latitude intervals of 10°. Since neither fF_2 nor longitude is a simple function of dip angle, the resultant fF_2/longitude curves show a variety of forms, some of which are reproduced in Fig. 1 in which alternate 10% sections are dotted. It is found that the change in fF_2 along a given line of latitude varies from about 20% to 100%.

Construction of Chart

The next step is to mark off on each curve the longitude limits which include the uppermost and lowermost 10% of the range covered by fF_2. In part of the Northern Hemisphere these two regions include the majority if not all of the fF_2 values. In the Southern Hemisphere, however, it was found that the rapid change in dip at some latitudes necessitated the insertion of several additional 10% regions besides those already mentioned.

In general the highest 10% regions occur over Europe, Asia and South America, while the lowest 10% lies over North America and the Indian Ocean. It is important to remember that within a zone as defined, the way in which fF_2 changes with longitude does not matter and consequently the same zone may include, at different latitudes, both the top and bottom 10% portions of the fF_2 range.

Having divided up each fF_2/longitude curve into 10% regions, the boundary points can be plotted on a graticule and joined up so as to form a series of closed zones (Fig. 2). In each of these zones, the change in fF_2 along a line of latitude does not exceed 10%. The graticule has been constructed in such a way that the length

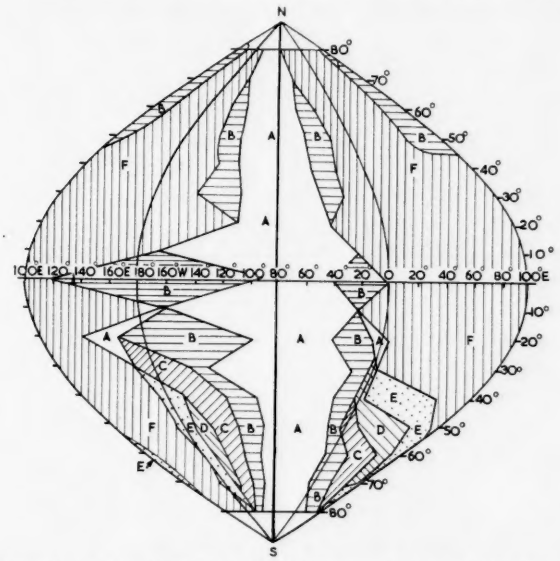


Fig. 2. Zones within which the change in fF_2 along any line of latitude does not exceed 10% (noon, September 1948).

of each latitude line is proportional to its true length and areas both at the poles and the equator are directly comparable. (Sinusoidal equal area projection.)

Although this map refers to noon at an equinox near sunspot maximum, an examination of a noon fF_2/dip-angle curve for sunspot minimum, and a midnight curve near sunspot maximum shows that the corresponding zone boundaries would be similar to those shown. At the solstices, the fF_2/dip-angle curve is assymmetrical about the zero dip line; the relation in the summer hemisphere is similar to that at the equinox but in the winter hemisphere the change in fF_2 with dip angle appears to be less rapid [5]. At the solistices therefore the zone system shown on Fig. 2 would become simpler in the winter hemisphere and would remain approximately as shown in the summer one.

Defects of Zoning Systems

Fig. 3 shows the boundaries of the existing E-, I- and W-zones and also those proposed by Oboril and Rawer. It is obvious that the W- and E-zones correspond roughly to the zones marked A and F respectively on Fig. 2. On the other hand the European

VOL. 2 1952

I-zone includes an area which could, without serious error, be included in the *E*-zone and this has been done in the modified zone system proposed by OBORIL and RAWER.

Although various modifications to the present zone system could be devised, in view of the fact that the zone boundaries in Fig. 2 do not correspond to any easily defined limits, it appears to be futile to continue the search for the "best" positions for the conventional *E*-, *I*- and *W*-zone boundaries. The possibility of defining the zones in terms of the boundaries shown in Fig. 2 is impractical, not only account of their unwieldy shapes, but also because the shape and the very

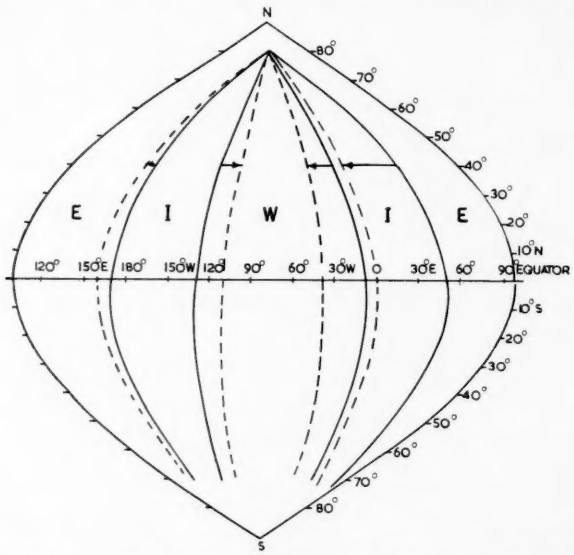


Fig. 3. Boundaries of *E*-, *I*- and *W*-zones (full lines). Modifications proposed by OBORIL and RAWER (chain lines).

existence of some of them would change from month to month. A further insurmountable difficulty is that in the Southern Hemisphere the changes in fF_2 are so great that if only three zones were defined as at present, then the *I*-zones would have to include changes in fF_2 amounting to 30–40% in some areas.

SUGGESTIONS FOR FUTURE PRACTICE

From what has been said it is suggested that to approach the question of representing the longitude effect in F_2-region by way of geographical zones is unsound and that no satisfactory solution can ever be found in this way. The possibility of creating a series of circular zones, each centred on an important transmitter or receiver, is not an economical one. The total number of such zones would be large and in certain parts of the world they would overlap.

If it is admitted that the present three-zone system introduces errors which are intolerable then, apart from the apparently impractical possibility of defining additional zones, the only reasonable alternative [6] is to construct a set of charts each showing the world distribution of fF_2 for a given value of either L.M.T. or U.T. If, for example, such charts were available at intervals of two hours, then twelve

of them would be required to cover the twenty-four hours as compared with three under the present system. If L.M.T. charts were used, it is conceivable that by the judicious choice of the values of L.M.T., a smaller number would suffice without causing serious loss of accuracy.

From the commercial user's point of view a set of U.T. charts would appear to provide the best solution because the fF_2 values for both ends of any long-distance circuit could be obtained from a single chart; on the other hand, the use of L.M.T. charts would in general require the use of two charts for this purpose and an interpolation of fF_2 at one end to make the U.T.'s correspond at both ends.

Conclusions

Assuming that a relation between fF_2 and magnetic dip angle of the type discussed by Appleton is valid, it is shown that the longitude effect in fF_2-region cannot be represented satisfactorily by any system of geographical zones.

The only satisfactory solution to the problem appears to be the adoption of a set of charts covering the twenty-four hours at suitably chosen intervals and each chart referring either to a constant L.M.T. or U.T. The U.T. charts would appear to be preferable where their use in commercial organizations is envisaged.

The work described above was carried out as part of the programme of the *Radio Research Board*. This paper is published by permission of the Director of the *National Physical Laboratory* and the Director of Radio Research, of the *Department of Scientific and Industrial Research*.

VOL. 2 1952

Reference

[1] Appleton, E. V.; Two Anomalies in the Ionosphere, Nature, Lond. 1946 **157** 691. [2] Ionospheric Radio Propagation, National Bureau of Standards (Washington) Circular 462, 1948, p. 98. [3] Oboril, F. and Rawer, K.; The Longitude Effect in the F_2-Layer of the Ionosphere, C. R. Acad. Sci., Paris 1949 **228** 1962. [4] Appleton, E. V.; Studies of the F_2-Layer in the Ionosphere, J. Atmos. Terr. Phys. 1950 **1** 106. [5] Lépéchinsky, D. and Pillet, G.; The Influence of Terrestrial Magnetism on the Ionization of F_2 Layer, Notes Préliminaires du Laboratoire National de Radioélectricité, Paris, 1950, No. 139. [6] Appleton, E. V.; The Investigation and Forecasting of Ionospheric Conditions, J. Inst. Elect. Engrs 1947 **94** (Part IIIA) 186.

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 266 to 271. Pergamon Press Ltd., London

A modified McClelland method for measuring ionic mobilities

P. J. Nolan and P. J. Kenny

University College, Dublin

(Received 18 February 1952)

ABSTRACT

The limits of the large ion group spectrum in room air are investigated by a modification of the McClelland mobility method. An aggregation scheme based on these limits is suggested. This scheme accounts to some extent for the large ion grouping previously observed.

In the course of work on large ions in room air we have recently used a modification of the McClelland mobility method. The modified method possesses several interesting features which seem worthy of record. The method first used by McClelland [1] is well known and is usually attributed to Rutherford and Zeleny. There are many published accounts which include the unnecessary and unjustifiable assumption that the air velocity in the cylindrical condenser is independent of the distance from the axis of the cylinders.

Original McClelland Method

A steady current of air containing ions is drawn through a cylindrical condenser, the inner and outer radii of which are a and b. The length of the inner electrode is l and the potential difference between the electrodes is V. The radial field at a distance r from the axis is $V/r \log_e((b/a)$ and the radial velocity at that point of an ion of mobility k is $kV/r \log_e(b/a)$. If u is the air velocity at a distance r from the axis, taken as the axis of x, then the differential equation of the path of an ion is

$$u\,r\,dr \log_e(b/a) = k\,V\,d\,x.$$

We now write three equations with increasing values of V

$$\log_e(b/a) \int_a^y u\,r\,dr = kVl, \tag{1}$$

$$\log_e(b/a) \int_a^b u\,r\,dr = kVl, \tag{2}$$

$$\log_e(b/a) \int_a^b u\,r\,dr = kV\,x. \tag{3}$$

The saturation voltage is given by eq. (2). Since $2\pi \int_a^b u\,r\,dr = Q$, the volume of air per second we obtain the standard formula $Q \log_e(b/a) = 2\pi\,k\,Vl$ which may be written $Q = 4\pi\,k\,V\,C$ where C is the capacity of the inner electrode. Distortion of the field is taken into account by using the experimental value for this capacity.

If the value of the field is less than the saturation value the ions reaching the electrode come from a smaller area of the cross-section. The volume of air from which the ions are collected in unit time is $2\pi \int_a^y u\,r\,dr$ so that, provided the ionic concentration is uniform, the current is proportional to the applied field and Ohm's Law

holds [eq. (1)]. This gives the theory of the aspiration method of measuring the conductivity of the atmosphere.

If the value of V is greater than the saturation value, eq. (3) applies. The ions are collected on a length x less than l and the current remains constant.

Two distinct groups of ions will give a current-voltage curve of type $OABC$ of Fig. 1. The voltages for points A and B give the saturation voltages and the relative charge concentrations are given by OR and RS. The average mobility can be obtained from the voltage for the point M. With three or more groups corresponding considerations apply.

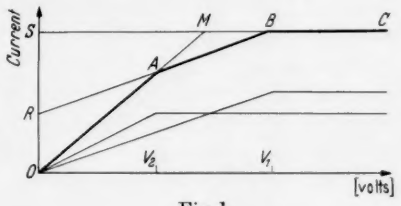


Fig. 1.

Modified Method

The modified McClelland method now described consists in using a fixed high voltage and measuring the current to the inner electrode for various air-currents through the tube. A graph of current against air-blast is plotted. If only one group is present the current-blast curve consists of two straight lines, one through the origin and one parallel to the blast axis. For low blasts the fixed voltage is sufficient to produce saturation and eq. (3) applies. The quantity of electricity collected is directly proportional to the blast. When the blast is increased, saturation ceases at the critical blast given by eq. (2). For larger values of the blast, eq. (1) applies and the electric current is constant as it is proportional to $\int_a^y u\,r\,dr$. As u varies y varies so as to keep this quantity constant. This region of the current-blast curve corresponds to the initial Ohm's Law or conductivity straight line of the ordinary McClelland method. The initial straight line of the current-blast curve corresponds to the saturation region of the ordinary method.

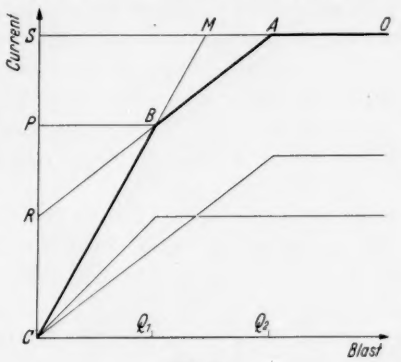


Fig. 2.

The current-blast curve of two groups is shown in Fig. 2. The critical blasts are given by points A and B; the relative contributions are given by RS/CQ_2 and CR/CQ_1 or since $RS/CQ_2 = RP/CQ_1$ the ratio of the contributions is RP/CR. If three or more groups are present the relative contributions are given by intercepts on CP by lines through the first corner B parallel to the various straight lines on the current-blast curve. The average mobility may be deduced from point M. The portions CB of Fig. 1 and 2 are the saturation regions and the parts OA the conductivity regions.

Comparison of Two Methods

With the current-voltage method it is necessary in ensuring that the lowest mobility present has been detected, to adjust the blast so that the point B, Fig. 1 occurs well below the highest voltage used. In other words, the line BC Fig. 1 must extend over a considerable voltage range. In the variable blast method the highest voltage may be used as the fixed voltage.

VOL. 2 1952

One of the great difficulties in experiments with natural ions is that the quantity of electricity is small with consequent trouble caused by drifts, leaks and unsteadiness in the high voltage supply. With the variable blast method, full use is made of the highest suitable voltage for all groups so that the currents measured are much larger. This is especially the case with the smaller ions. The method is particularly useful for detecting the smallest ion present.

The diagrams in Figs.1 and 2 have been drawn to illustrate these features. Two groups of mobility ratio 2/1 and concentration ratio 3/4 give in Fig. 1 a saturation current of 7 units. The fixed voltage of Fig. 2 is that of point C of Fig. 1 and is 3/2 the saturation voltage. As a result the blast at point B of Fig. 2 is 3/2 the fixed blast of Fig. 1 and the contribution of the slower group is 6 units instead of 4 units. As the fixed voltage of Fig. 2 is three times that of point A of Fig. 1 the contribution of the faster group is 9 units instead of 3 units. The maximum current in the variable blast method is thus 15 units in contrast with 7 units of the other method. We have taken a simple case as example; in practice the current magnification is larger as the ratio of the extreme mobilities if of the order 30.

One of the advantages of the variable blast method is that the high voltage applied to the outer tube is not changed during an experiment. This increases the probability of steadiness in the voltage and entails constancy in drift corrections.

The time which the ions spend in the field is much shorter in the variable blast method. Errors caused by diffusion and mutual repulsion are as a consequence reduced. This consideration applies more to small ions than to intermediate or large ions.

The variable blast method cannot obviously be applied to the examination of artificial ions in cases in which the time between formation and measurement must be constant. With natural ions and long inlet tubes allowance may have to be made for diffusion with slow blasts and for turbulence with fast blasts.

Zeleny Method with Variable Blast

Resolution of groups in the Zeleny method depends on the ratio of the total length of the central electrode to the length of the insulated portion. It is desirable that this ratio should at least be larger than the ratio of adjacent mobilities. In Fig. 3 the current-blast curve for two groups with a Zeleny tube is given. The ratio total length to insulated length is 3/1; the mobility ratio is 2/1 and the concentration ratio 3/4. Mobilities can be deduced from points X and B and from points Y and A. The maximum current is one-third the maximum current of Fig. 2. In the Zeleny method with variable blast a stepped curve will be obtained if the ratio total length to insulated length is greater than the ratio of adjacent mobilities.

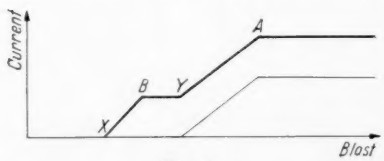


Fig. 3.

Large Ions in Room Air

The modified McClelland method as outlined was applied to the examination of large ions in room air in University College, Dublin. The dimensions of the apparatus were as follows:

Length of inner electrode = 80·5 cm

External diameter of inner electrode = 1·90 cm

Internal diameter of outer electrode = 2·91 cm
Measured capacity of inner electrode = 98·6 cm.

Small ions were removed by a short ion tube with low voltage at entrance.

The air-blast was obtained from a rotary pump driven by an electric motor. Blasts of from 2 to 60 l per min were used. Two rotameters of different ranges measured the air-blasts. The current to the inner electrode was determined by a balance arrangement similar to that used by J. J. NOLAN and P. J. NOLAN [2]. A LINDEMANN electrometer was employed as a null instrument. During a measurement the needle of the electrometer was maintained at zero by adjusting the potential of the outer electrode.

Owing to fluctuations it is impossible to perform mobility measurements on outside air with one ion counter. With room air satisfactory results can be obtained but it is important that conditions should be steady before beginning the experiment proper. Draughts, the opening and closing of doors and all such disturbing influences must be eliminated. It was found advisable to allow about ten minutes to elapse from the time of starting the rotary pump in order that the air circulation caused by it should become regular.

VOL. 2 1952

It was not our intention to catalogue all the groups in room air. Our purpose was to apply the variable blast method to the determination of the lowest and the highest mobility present. As we have seen the method is particularly suited for the latter determination. Using a fixed field of 600 v we made six satisfactory measurement of the lowest mobility, the values lying between $1{\cdot}0 \times 10^{-4}$ cm/sec/v/cm and $1{\cdot}8 \times 10^{-4}$. This ion group represented about 30 % of the total ionisation. We consider that ion groups of lower mobility would have escaped detection if present up to say 5 % of the total ionisation. BOYLAN [3] examined the mobility spectrum of atmospheric large ions paying particular attention to the question of the largest ion. In addition to an ion of mobility $1{\cdot}3 \times 10^{-4}$ he found ions in the range $0{\cdot}9 \times 10^{-4}$—$0{\cdot}5 \times 10^{-4}$ corresponding to ions found in artificial ionisation by McCLELLAND and NOLAN [4]. NOLAN and DEIGNAN [5] using the nucleus-field method which is not suited for detecting ion groups in small concentration found that the limiting mobility in air stored for one or two hours was $1{\cdot}0 \times 10^{-4}$. Nuclei stored for 45 hrs gave as the limiting mobility $0{\cdot}35 \times 10^{-4}$, corresponding roughly to BOYLAN'S minimum value.

In the determination of the largest mobility a fixed voltage of 120 volts was used. The largest mobility was found to be 41×10^{-4} and the group represented about 5 % of the total ionisation. On two occasions there were indications of the presence of more mobile ions; on one occasion the limiting mobility was 18×10^{-4}. Our result agrees with the previous Dublin observation [6] that in room air the ion of mobility 43×10^{-4} is the fastest ion present in any considerable quantity among the large ion group. This does not exclude the possibility of the presence of intermediate ions of mobility about 500×10^{-4} in small concentration. In room air in Dublin the large ion concentration is of the order 10 000 and the small concentration of the order 200. The fastest group found by NOLAN and DEIGNAN in air stored for one or two hours was 6×10^{-4}; but as we have stated their method is not suitable for the detection of groups in small concentration and the faster ions disappear by coagulation during storage. HOGG [7] found in atmospheric air at Kew that the intermediate ion spectrum extended to about 150×10^{-4}.

Existence of Distinct Groups of Ions

The production of groups of ions by spraying water, by bubbling air through liquids and by passing air over phosphorus has been observed by J. J. Nolan [8] and by J. A. McClelland and P. J. Nolan [4]. The existence of such groups of ions has been questioned by Blackwood [9] and by Busse [10]. Blackwood using the McClelland method did obtain breaks in the slope in the current-voltage curve with ions produced by spraying water. He worked over 59 of these in detail and decided that the mobilities did not fit into any group system. The mobility values are recorded in his paper. Omitting a few outlying values there are 52 mobilities lying between 6×10^{-4} and 170×10^{-4}. It is interesting to note that despite Blackwood's failure to find evidence of grouping 12 of these mobilities are in the range 35×10^{-4} to 45×10^{-4}, constituting apparently a group corresponding to the largest mobility in room air.

Distinct groups similar to those produced artificially were found by Boylan in atmospheric air. The existence of the groups produced by spraying was confirmed by Nolan and O'Keeffe [11]. The mobility of the ions in the corona discharge was investigated by Young [12] and a large number of distinct groups was observed. It is possible that these groups are not formed in the same way as the previous groups. The corona ions may be solid particles; it is probable that all the other ions are liquid. The production of discrete groups by spraying liquids was confirmed by Chapman [13].

Hogg showed that the intermediate ions of the atmosphere exist in separate groups with distinct mobilities and that the variation of these mobilities with humidity indicate that the ions are composed of some hygroscopic substance probably sulphuric acid. He further showed that in the range investigated there are 15 distinct group sizes of volume ratios 1, 2, 3 . . . 15, resulting from agglomeration. Vassails [14] has examined the nuclei produced by ultra-violet light in air. The mobilities of the charged nuclei form a discontinuous series often linked by simple ratios (2/1, 3/2, 4/3). He points out that this result suggests that distinct groups are due to multiple charges of ions of the same size. This suggestion has previously been made (Israël and Schulz [15]). The experiments of Vassails are noteworthy in that distinct mobilities are obtained by a statical method, that is a method not involving air flow. In a review of the question of ion groups in "Conduction of Electricity through gases" (J. J. Thomson and G. P. Thomson, 1928) the verdict is against the reality of groups; it is suggested that the observation of groups is associated with eddies in the air-flow. In our opinion the exact opposite is to be expected; the mixing produced by turbulence would tend to obliterate evidence of any groups present. In the variable blast method however, the onset of turbulent flow might give a kink in the current-blast curve and a second kink might occur at the "slip" velocity (Dowling [16]).

It is obvious from the investigations quoted that the existence of discrete groups of large and intermediate ions is beyond question. On the part played by size and charge in producing the various mobilities we are able produce some fresh evidence. We assume in spite of some striking evidence to the contrary [11] that the mobility is proportional to the charge. The experiments described earlier were carried out during an investigation on the average charge on large ions in room air. These results are not yet published. The median value of 40 determinations of the average charge was 1·67 electronic charges; this average charge showed no marked dependence on

mobility. As the ratio of the limiting mobilities 41×10^{-4} and $1{\cdot}5 \times 10^{-4}$ is 27 the theory of a single size is excluded.

Using (1) the equation $k = 40\,D$ applicable to singly-charged ions where k the mobility is expressed in cm/sec/v/cm and D is the diffusion coefficient and (2) the Stokes-Cunningham-Millikan relation between D and the radius we have calculated the radii corresponding to mobilities $41{\cdot}6 \times 10^{-4}$ and $1{\cdot}52 \times 10^{-4}$. The radii $1{\cdot}13 \times 10^{-6}$ cm and $7{\cdot}05 \times 10^{-6}$ give a volume ratio 3^5. In Table 1 we have inserted between these limits mobilities corresponding to volumes in geometrical progression with a ratio of three. We have also extended the series to ultra large ions of mobilities $8{\cdot}5 \times 10^{-5}$ and $5{\cdot}0 \times 10^{-5}$ which have been found in air passed over phosphorus and in atmospheric air. In this scheme the Langevin ion of mobility $3{\cdot}3 \times 10^{-4}$ appears with mobility $2{\cdot}8 \times 10^{-4}$. However in view of Hogg's conclusions with regard to the effect of relative humidity we cannot expect exact agreement. The scheme gives roughly the 2/1 relation between adjacent mobilities found in this region for prominent ion-groups especially in the alcohol and phosphorus experiments.

Table 1

$10^3 k$	$10^6 D$	$10^6 r$	*Volume*
4·16	104	1·13	1
2·06	51·5	1·63	3
1·04	26·0	2·35	3^2
0·53	13·2	3·39	3^3
0·28	7·0	4·89	3^4
0·152	3·8	7·05	3^5
0·085	2·12	10·2	3^6
0·050	1·25	14·7	3^7

It must be admitted that the value 1·67 which we have obtained for the average charge raises several difficulties; we do not intend to pursue them at present. One difficulty for example is that the room nuclei investigated have an average radius of about 3×10^{-6} whereas the critical radius at which double-charging begins is $1{\cdot}5 \times 10^{-6}$ (Bricard [17]) or 4×10^{-6} (Nolan and Kennan [18]). From Bricard's figures we deduce that the average charge for radius 10×10^{-6} is 1·47; Nolan and Kennan find for the same radius an average charge 1·29.

If we ignore this difficulty and assume that a considerable number of the smaller nuclei have double charges we may attribute the prominence of the groups of Table 1 to the 2/1 ratio between neighbouring mobilities. This consideration applies especially to the groups in the upper portion of the Table. The suggested volume ratio 3 may be associated with the fact that a fraction of the nuclear coagulations, those between large ions of opposite sign, entail the disappearance of an ion. These two considerations may in the mobility region under review account for the absence of evidence of intermediate aggregations. It is noteworthy that although Hogg's scheme is a 1, 2, 3, 4, etc. type of aggregation the groups in any particular experiment had a wider separation.

References

[1] McClelland, J. A.; Phil. Mag. 1898 **46** 29. [2] Nolan, J. J. and Nolan, P. J.; Proc. Roy. Irish Acad. 1935 **42** 15. [3] Boylan, R. K.; Proc. Roy. Irish Acad. 1931 **40** 76. [4] McClelland, J. A. and Nolan, P. J.; Proc. Roy. Irish Acad. 1916 **33** 24; 1918 **34** 51; 1919 **35** 1. [5] Nolan, P. J. and Deignan, J.; Proc. Roy. Irish Acad. 1948 **51** 239. [6] Nolan, J. J. and de Sachy, G. P.; Proc. Roy. Irish Acad. 1927 **37** 71. [7] Hogg, A. R.; Proc. Phys. Soc. 1939 **51** 1014. [8] Nolan, J. J.; Proc. Roy. Irish Acad. 1916 **33** 9. [9] Blackwood, O.; Phys. Rev. 1920 **16** 85. [10] Busse, W.; Ann. der Phys. 1925 **76** 493; 1927 **82** 697. [11] Nolan, J. J. and O'Keeffe, J. G.; Proc. Roy. Irish Acad. 1933 **41** 25. [12] Young, W. M.; Phys. Rev. 1926 **28** 129. [13] Chapman, S.; Phys. Rev. 1937 **52** 184. [14] Vassails, G.; Ann. de Phys. 1949 **4** 125. [15] Israël, H. and Schulz, L.; Terr. Mag. 1933 **38** 285. [16] Dowling, J. J.; Proc. Roy. Dub. Soc. 1912 **13** 375. [17] Bricard, J.; Jour. Geophys. Research 1949 **54** 39. [18] Nolan, P. J. and Kennan, E. L.; Proc. Roy. Irish Acad. 1949 **52** 171.

VOL. 2 1952

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 272 to 281. Pergamon Press Ltd. London

A search for radio echoes of long delay

K. G. BUDDEN* and G. G. YATES**

(Received 20 February 1952)

ABSTRACT

A description is given of some attempts to detect long delayed radio echoes of the type reported by VAN DER POL [2], STØRMER [1] and others, with delays between 3 sec and 15 sec. It has been suggested that such echoes might arise by reflection from ionised corpuscular streams from the sun. To investigate this, frequencies were used which would penetrate the ionosphere at most times of day, and it was also necessary to use the highest possible transmitter power. Both these requirements are fulfilled by some commercial short wave transmitters, and we are indebted to *Messrs. Cable and Wireless Ltd.*, who arranged special transmissions from Ongar.

No echoes of long delay were observed. Owing to the large number of short wave commercial transmitters now operating, the interference level was probably far higher than it would have been in 1928. Nevertheless, it is possible to estimate the minimum signal power which could be detected by the receiving apparatus. The probable size and density of the clouds of ions emitted from the sun is then examined, and it is concluded that such clouds are not likely to be dense enough to give detectable radio echoes. It is tentatively suggested that the echoes of long delay heard by previous observers were associated in some way with ionised regions fixed relative to the earth, and that the failure to observe echoes in the present investigation was due to the choice of frequencies which were too high. Some suggestions are made for the planning of future attempts to observe long delayed echoes.

1. INTRODUCTION

Several reports have been made of radio signals received from transmitters on the earth after a time delay of several seconds, (for example, STØRMER [1], VAN DER POL [2]). The delay varies between about 3 and 15 sec, and the echoes are usually very distorted. No completely satisfactory explanation has been put forward, but various suggestions have been made. VAN DER POL [2] and FUCHS [3], [4] have suggested that the echoes may be associated with the ionosphere, and have described mechanisms by which the long delays might occur. STØRMER [1], [5] considered that the echoes might arise by reflection from clouds of ions at great distances from the earth, and PEDERSEN [6] has shown how radio waves might be guided over long paths well outside the earth by "belts" or "bands" of ions.

There is now considerable evidence that ionised clouds are sometimes emitted from the sun towards the earth, and give rise to terrestrial magnetic storms. The experiments described in this paper were planned to investigate whether reflections of radio waves could be obtained from clouds of ions well outside the earth. It was therefore decided to use frequencies which would penetrate the ionosphere, and it was also obviously desirable to use the highest available transmitter power. These requirements are both admirably fulfilled by commercial short wave transmitters. We are greatly indebted to *Messrs. Cable and Wireless, Ltd.*, who made available two of their transmitters at Ongar, Essex. The exact frequencies used depended mainly on which transmitters could be released from commercial service. The experiments

* Cavendish Laboratory, Cambridge.
** Solar Physics Observatory, Cambridge.

were made with the transmitters GMC, frequency 13·455 Mc/s, power 30 kw, and GMV, frequency 20·675 Mc/s, power 9 kw. A description of the aerial systems and receivers is given in section 4.

No long delayed echoes were observed in any of the experiments. Observations were sometimes difficult to make owing to heavy interference from other commercial short wave transmitters. These were much more numerous than in 1928 and 1929 when STØRMER and VAN DER POL made their observations. The sensitivity of the apparatus is examined in sections 6, 7, 8 and 9, and estimates are made of the density and size of the clouds of ions that would be necessary to give the minimum detectable signal. There is very little information regarding the density of the clouds emitted from the sun, but the suggestions that have been made are all very much lower than would be required to give detectable radio echoes. We conclude that the long delayed echoes observed in the past were probably associated with some ionised region close to the earth, and the evidence in favour of this is summarised in section 10. Some suggestions for the planning of future attempts to observe long delayed echoes are made in section 11.

2. Previous Observation of Echoes of Long Delay

The most striking observation of echoes of long delay were probably those of VAN DER POL [2], and STØRMER [1], with signals of frequency 9·55 Mc/s from the transmitter PCJJ at Hilversum. The echoes were observed on several different days in March and October, 1928. The delays between the times of arrival of the main signal and the echo varied between three seconds and fifteen seconds, and the echo signal was usually very distorted. The echoes were again observed in 1934 during a series of tests organised by the *World Radio Relay League*, and inaugurated by Sir EDWARD APPLETON. The observers were British radio amateurs, and the frequencies used were 9·5 Mc/s, and 6·7 Mc/s. The echoes had characteristics very similar to those reported by STØRMER and VAN DER POL. Similar echoes have also been reported by GALLE, TALON and FERRIÉ [8].

3. Possible Explanations

An explanation of the occurrence of these echoes was put forward by VAN DER POL in his original communication to "Nature," [2]. He suggested that they might be due to a long transit time in a region of the ionosphere where the group velocity is very low. Such a region is well known to occur at the level in the ionosphere where the refractive index of the medium is close to zero. APPLETON [9] pointed out that this explanation is probably untenable, because the waves would be attenuated in the ionosphere to an extent which can be estimated. He showed that after times of the order of the observed delays, the signal amplitudes would be far too small to be observed at all. The explanation would only be a possible one if the delay occurred at a level in the ionosphere far higher than the known levels of the E and F regions, where the frequency of collisions of electrons with gas molecules is very small. This explanation has been examined in more detail by PEDERSEN [6], BREIT [10], and by FUCHS [3] ,[4], who deduced that the electron collision frequency in the region where the delay occurs must be of the order of 0·5 sec^{-1}. THOMAS [11] estimated that even if the gas pressure is so low that collisions of the electrons with gas

molecules are negligible, there is still an effective collision frequency of the order of 10^4 sec^{-1} due to the electrostatic forces between the electrons. Hence, the signals would be attenuated far too heavily to give observable echoes. Joos [12] also examined the suggestion that the observed delay arose because of a low group velocity somewhere in the path of the waves. He concluded that, quite apart from possible absorption, the dispersion which is necessarily associated with a low group velocity would be so great that it would reduce the signal amplitude far below the minimum detectable level.

A second possible explanation is that the echoes are due to the reflection of radio waves from clouds of ions or electrons at some distance from the earth. This explanation was favoured at first by Størmer [1] [13], who suggested that streams of ionised particles shot from the sun towards the earth may be responsible for the echoes. Such streams of ions moving towards the earth would have their trajectories modified by the earth's magnetic field, and Størmer [13] has shown that they would be distributed in a region bounded on one side by a toroidal surface surrounding the earth. The particles would be excluded from the inside of this toroid, and he suggested that the echoes might be evidence of the reflection of radio waves from its inner surface. The estimated size of the toroid would account for delays of the right order (7 sec).

A third explanation was given by Pedersen [6], who showed how radio waves could be guided along "belts" or "bands" of ions. He suggested that such guided waves could travel over large curved paths starting and ending near the earth. Størmer [14] discussed this suggestion, and reviewed the possibility that such paths were associated with a ring of current round the earth, due to the motion of charged corpuscles. This ring would be in the earth's magnetic equatorial plane, and Størmer showed that, with this explanation, the conditions for receiving long delayed echoes are most favourable when the line from the earth to the sun is also in this plane. He gives data which show that echoes were in fact observed most frequently at these times.

Explanations in terms of many reflections in the ionosphere have been suggested by Janco [15] and Jelstrup [16]. An explanation based on guided waves was suggested by Jonescu and Mihul [17]. Appleton [9] suggested that the waves might be repeatedly focussed at the antipodes and near the transmitter, so as to give an audible signal after several seconds.

The suggestions of Størmer and Pedersen have acquired a new interest in view of recent theories to explain the occurrence of terrestrial magnetic storms. It is known that magnetic storms frequently start at times of the order of 20 hrs, after the occurrence of a bright solar eruption. It has therefore been suggested that simultaneously with the occurrence of the eruption, a burst of particles is shot out from the sun, and takes about 20 hrs to reach the earth. There is ample evidence in favour of this theory, (see, for example, Chapman [18]) and it was therefore of special interest to investigate whether long delayed echoes could be heard during the period immediately following a solar flare.

Størmer's conclusions were based on the calculated trajectories of single charged particles in the earth's magnetic field. More recently Chapman and Ferraro [19], [20], [21] have shown that the clouds of particles moving from the sun towards the earth are most likely to be electrically neutral, ionised clouds, since clouds of charges

of one sign would be rapidly dispersed by electrostatic repulsion. The effect on a moving, electrically neutral cloud, of the earth's magnetic field has been discussed by CHAPMAN and FERRARO. They showed that the particles are deflected in such a way that in the cloud there is a hollow surrounding the earth. This hollow is somewhat similar to that pictured by STØRMER. At a later stage the streams of ions move into a closed structure surrounding the earth in the form of a current ring, with which is associated the start of the terrestrial magnetic storm. This is similar to the ring described by STØRMER [5]. It is of interest to consider whether the observed long delayed echoes could originate by reflection at the inner surface of the hollow, or from the current ring. This is considered further in sections 8 and 9, and we conclude that such reflections would be far too weak to give echoes. A similar conclusion was reached by DOSTAL [22], who also considered whether the echo amplitude could be increased by a focussing mechanism due to the concavity of the reflecting surfaces.

4. EXPERIMENTAL DETAILS

Attempts to observe echoes of long delay were made at Cambridge during 1947, 1948, and 1949. When the experiments were planned, their main interest was thought to be the possible detection of reflecting or scattering objects at great distances from the earth. Fairly high frequencies were therefore used so that waves travelling radially outwards from the earth could penetrate the ionosphere. Apart from this, the choice of frequencies was dictated mainly by the availability of the transmitters, which had to be diverted from commercial service for the special transmissions. Two frequencies were used, namely 13·455 Mc/s, (power 30 kw), and 20·675 Mc/s, (power 9 kw). The transmissions were provided by Messrs. *Cable and Wireless, Ltd.*, from their transmitting station at Ongar, Essex.

VOL. 2 1952

The transmitting aerial used for both frequencies was a horizontal dipole approximately half a wavelength long, at about one quarter wavelength above the ground. This arrangement was chosen so as to give maximum radiation in a direction radially outwards from the earth. Unmodulated continuous waves were used, keyed on and off with Morse characters. During the first quarter hour of each transmission, the transmitter sent its call sign and V's continuously, so that the receivers could be tuned. It then sent single call signs at approximately twenty second intervals. In some transmissions the call sign and V's were sent for five minute periods every half hour, so that the receiver adjustments could be checked. Most of the transmissions lasted two hours.

The receiver had two stages of radio frequency amplification with V R 91 valves. These were followed by a mixer valve fed from a local oscillator, which was crystal controlled. The output of the mixer was fed to an I.F. amplifier operating at 455 Kc/s with a band width of about 100 Kc/s. This was followed by a second mixer valve, fed from a second local oscillator, which could be manually adjusted. The output of the second mixer was fed to a second I.F. amplifier, operating at 130 Kc/s, with a band width of about 2 Kc/s. The output of this was caused to beat with a continuous oscillator, so as to give a note of about 1 000 c/s. This was fed to an audio frequency amplifier incorporating a narrow band filter, and thence to a cathode follower valve. The output of this was fed to a pair of headphones, and to the vertically deflecting plates of a cathode ray oscillograph. The overall bandwidth of the

receiver was about 200 c/s, and was determined by the narrow band filter in the audio frequency amplifier.

If an observed echo were returned from a moving object, it would suffer a Doppler change of frequency. An additional unit was therefore added, in later experiments, to allow a search to made on other nearby frequencies. It was fed from the output of the first I.F. amplifier, and consisted of a local oscillator, an I.F. amplifier at 130 Kc/s, and output circuits similar to those in the main receiver. Its output was connected to the vertically deflecting plates of a second cathode ray oscillograph. The two parts of the receiving apparatus will be referred to as the "main receiver", and the "Doppler receiver". The range of frequencies to which the Doppler receiver could be tuned was determined by the bandwidth of the first I.F. amplifier, and was about 100 kc/s on the high frequency side of the transmitter frequency. A frequency change of this amount would be produced by a reflecting or scattering object having a "line of sight" velocity of about 1100 km/sec towards the earth.

The horizontally deflecting plates of the cathode ray oscillographs were connected to a source of time-base voltage which gave a slow linear sweep to the spot, with a duration of about 20 sec, and triggered by the main signal. The screens of the cathode ray oscillograph were photographed using paper moving slowly and continuously, so that successive traverses of the time-base appeared as parallel lines about 7 cm long and 1 cm apart.

In the tests on 13·455 Mc/s, the receiving aerial was a horizontal half wave dipole at about one quarter wavelength above the ground. In the tests on 20·675 Mc/s, a few observations were made with this type of aerial, but in most of the experiments the aerial used was a vertical Vee aerial, designed to have maximum sensitivity for radiation coming radially inwards to the earth. The polar diagram of this aerial was tested, using a small oscillator in a captive balloon. The diagram was found to have the shape expected, and the aerial had a power gain of approximately 8 to 9.

5. Method of Observation and Results

During a transmission, as soon as the transmitter began sending call signs at 20 sec intervals, the camera was switched on, and an observer listened to the received signals. There was usually only one observer, and he could listen to the output either of the main receiver, or of the Doppler receiver, but not to both. As far as possible, notes were made of any signals received, such as atmospherics, signals from other transmitters, or other man-made interference. The volume of traffic in the frequency band near 13·455 Mc/s was often heavy, and distant transmitters working practically on the same frequency sometimes caused trouble. Another effect of the heavy traffic was that the frequencies to which the Doppler receiver could be tuned were severely limited. The procedure with this receiver was to search near the transmitter frequency for a frequency band which sounded free from signals, and to leave the Doppler receiver tuned to this frequency. After five or ten minutes, the procedure was repeated, and if possible a new frequency was chosen. The displaced frequency used was always higher than the transmitter frequency, because it was thought that the echoes might most probably be due to streams of ions moving towards the earth. In view of the interference and the very limited number of frequencies that could be used, it is probable that the search for echoes with the Doppler receiver was of little value.

The heavy interference observed with the frequency 13·455 Mc/s meant that over 50 % of the signals examined were considered to be spoilt by interference. This was the main reason for changing to the frequency 20·675 Mc/s. At the higher frequency, interference was less, partly because there were fewer commercial transmitters in the frequency band, and partly because the frequency was above the maximum usable frequency (M.U.F.) for many transmission paths. Even so, the interference level was still high, and about 20 % of the signals examined were considered to be spoilt by interference. On the higher frequency, a very strong "round the world" echo was often observed (delay 0·13 sec).

At first, in a search for echoes which might occur at any time, transmissions were arranged at regular intervals, usually on one day in each fortnight. In each day transmissions were usually made during the intervals 0001–0200 hrs, 0600–0800 hrs, 1200–1400 hrs, and 1800–2000 hrs. In later experiments the transmissions were specially arranged when reports were received of solar flares or of other phenomena which might be associated with the emission of solar corpuscular streams. In these cases, four transmissions per day were made for a period of three days immediately following the occurrence of the flare. The majority of the observations were made in the years (1947–1949) near the epoch of maximum sunspots, and were spread fairly uniformly through the seasons.

VOL. 2 1952

About 100 transmissions, each using about 200 signals, were made on a frequency of 13·455 Mc/s. During 22 of these, both the main receiver and the Doppler receiver were used. 35 two-hour transmissions were made on a frequency of 20·675 Mc/s, during all of which both receivers were used. In all the experiments, the total number of signals emitted was about 27 000. In no case was any signal received which could be considered as an echo of long delay, or which resembled in any way the echoes described by previous observers.

6. The Sensitivity of the Apparatus

A theoretical lower limit to the sensitivity of the equipment was set by the thermal noise generated in the input circuit of the receiver with a power equal to $4kTB$, where B is the bandwidth, T is the absolute temperature, and k is Boltzmann's constant. With an overall bandwidth of 200 c/s, and the temperature of the input circuit 300° K, the theoretical thermal noise power was thus about $3{\cdot}2 \times 10^{-18}$ w. In actual tests with a signal generator, it was found that an input signal power of 10^{-16} w gave a signal-to-noise ratio of about 4:1 in amplitude at the output of the receiver. This corresponds to a noise power level of about $6{\cdot}2 \times 10^{-18}$ w. It was considered that a signal input of about 5×10^{-17} w was necessary to give a clear response on the photographic record, and that about 6×10^{-18} w would give an audible signal.

In the following discussion, it is assumed that the transmitting aerial had a power gain $G_t = 3$ in the vertically upward direction. The receiving aerial was assumed to have an effective cross-sectional area of $3\lambda^2/4\pi$, when intercepting radiation coming radially inwards to the earth, where λ is the wavelength. This is twice the value for a Hertzian dipole in free space. The values for a Hertzian dipole and a half wave dipole are roughly the same. A factor 2 is included to allow for reflection at the ground.

7. Detection of a Plane Perfect Reflector

If the radiation from the transmitter were reflected at a plane perfectly reflecting surface at a distance R from the earth, the power returned to the receiver would be

$$\{W G_t/4\pi(2R)^2\} \times \{3\lambda^2/4\pi\}. \quad (1)$$

At the greatest distance for detection this would be 10^{-17} w. With $W = 30$ kw, $\lambda = 22$ m, this would correspond to a distance $R = 1{\cdot}5 \times 10^8$ km. This would give a time delay between the main signal and an echo of 1000 sec. Thus a plane perfect reflector at a distance slightly greater than that of the sun would give the minimum detectable signal.

8. Sharply Bounded Ionised Region

The only way in which a cloud of ions in space could have a reflection coefficient approaching unity is that it should contain a concentration of electrons higher than the critical concentration for the radio frequency used. This critical concentration is given by

$$N_0 = \varepsilon_0 m p^2/4\pi e^2$$

where e, m = charge and mass of the electron,

ε_0 = permittivity of free space,

p = angular frequency of the wave.

For waves of frequency 13·5 Mc/s, this gives $N_0 = 2{\cdot}3 \times 10^6$ cm^{-3}.

It is apparently very difficult to estimate the density of particles in the streams of corpuscles travelling from the sun to the earth. Chapman and Ferraro [20], [21] tentatively suggest a figure of 100 to 1000 electrons per cc, when the stream is near the earth. Chapman and Bartels [23] mention a possible figure of 10^5 particles per cc in the stream when it leaves the sun. Clouds with densities greater than this, and with temperatures of the order of that of the sun's chromosphere would probably diffuse outwards rapidly. It therefore seems unlikely that electron densities as high as 10^6 cm^{-3} would occur in the streams of corpuscles, and so we should expect the stream to have a reflection coefficient very much less than unity.

We now consider what signals we should expect to be returned to the earth from clouds having electron densities of the order of 1000 cm^{-3}. These signals would depend upon the shape and size of the cloud, which is difficult to assess, but some idea of the upper limit of the reflecting power may be obtained by supposing that the cloud has a sharp plane boundary. We therefore consider a homogeneous sharply bounded medium, having an electron density $N = 1000$ cm^{-3}. The reflection coefficient for waves incident normally on the boundary is $\varrho = (\mu - 1)/(\mu + 1)$ where μ is the refractive index of the medium given by $\mu^2 = 1 - 4\pi N e^2/\varepsilon_0 m p^2$. Since μ is very nearly unity, we have with good accuracy $|\varrho| = \pi N e^2/\varepsilon_0 m p^2$. This reflection coefficient expresses the ratio of the amplitudes of a reflected wave to an incident wave. The reflection coefficient for power is the square of this, that is, $|\varrho|^2 = \pi^2 N^2 e^4/\varepsilon_0^2 m^2 p^4$. We must now multiply the expression (1) by this quantity to get the power returned to the receiver. If, as before, we equate the result to 10^{-17} w, then the resulting value of R is $6{\cdot}3 \times 10^4$ km. This distance would give a time delay between the main signal and an echo of only 0·42 sec, which is much

less than the longest delay times (of the order of 15 sec), reported by STØRMER and VAN DER POL. By a slight extension of the above calculation, we find that the electron density required to give a detectable echo with a delay of 15 sec, is $2{\cdot}7 \times 10^4 \text{cm}^{-3}$. But it should be remembered that the sharply bounded model gives only an upper limit to the reflecting power, and that diffuse clouds would return very much less energy to the earth. Hence, to obtain echoes with delays as great as 15 sec, electron densities very much greater than $2{\cdot}7 \times 10^4 \text{ cm}^{-3}$ would be required.

9. Scattering by Clouds of Electrons

As a further check on the above calculations, we may discuss how energy could be returned to the earth by scattering from clouds of electrons. CHAPMAN and FERRARO [19], [20], [21] suggest that the clouds of ions responsible for magnetic storms are very much larger than the earth, and eventually envelop it. A cloud would therefore subtend a large solid angle at the earth. Consider a part of a cloud of electrons which is at a distance R from the earth. If such a cloud were effective in giving an audible echo, the thickness of the region responsible would be much less than the distance travelled by radio waves in the duration of one Morse sign, say $\frac{1}{5}$ sec. We may take this as an upper limit for the thickness of that part of the cloud that is effective at any moment. We therefore consider a disk shaped region of the cloud on the surface of a sphere of radius R concentric with the earth. The thickness of the cloud is $d = 3 \times 10^4 \text{km}$, and its area is πR^2, so that its volume is $\pi R^2 d$, and it contains $\pi R^2 d \times N$ electrons, where N is the average electron density in the cloud. Now it is well known that the total power scattered by a free electron is $p = 8\pi e^4 P_0 / 3\varepsilon_0^2 c^4 m^2$, where c is the velocity of light, and P_0 is the flux of power in the incident wave. Since the electrons are assumed to be distributed randomly in the cloud, the powers in all the scattered waves must be added. Substitution of $P_0 = W G_t / 4\pi R^2$ then shows that the scattered power entering the receiver is

$$(W G_t / 4\pi R^2) \times (8\pi e^4 / 3\varepsilon_0^2 c^4 m^2) \times (3\lambda^2 / 4\pi) \times (1/4\pi R^2) \times \pi R^2 N d .$$

If $N = 1\,000 \text{ cm}^{-3}$, $R = 3 \times 10^5 \text{km}$ (corresponding to a delay of 2 sec), $W = 30 \text{kw}$, $G_t = 3$, then the received power would be $1{\cdot}0 \times 10^{-30} \text{w}$. This is smaller than the minimum detectable signal by a factor of 10^{12}.

It therefore seems reasonable to conclude that ionised clouds could not be detected by the scattering of radio waves transmitted from the earth, even if they had electron densities very much greater than those suggested by CHAPMAN and FERRARO. We must not, however, exclude the possibility that such clouds might be detected in the future due to their absorption of radio waves passing through them. For this, it would of course be necessary to use some extra-terrestrial source of radio waves, such as the sun or the stars.

There are reasons for believing that the radio echoes of long delay observed in the past are not connected with objects very distant from the earth. These reasons are summarised in the next section.

10. The Origin of Echoes of Long Delay

There appear to be three good reasons for supposing that echoes of long delay are associated with some system which is fixed in space relative to the earth. We discuss these reasons in turn.

VOL. 2 1952

1. There is no detectable difference between the frequencies of the echo signal and the transmitted signal. This was reported very strikingly by van der Pol [2]. We may assume that he was listening to a note whose frequency was of the order of 1000 c/s, and that he would readily have detected a change of pitch of less than a whole tone. This would mean a frequency difference of about 100 cycle sec^{-1}. If such a change of frequency were produced by reflection from a moving object, its velocity would be 1·5 km sec^{-1}. In fact, a velocity of 0·5 km sec^{-1} would probably have been detectable. It is difficult to see how the motion of particles travelling from the sun to the earth with velocities of the order of 1000 km sec^{-1} could have their motions modified so that their velocity relative to the earth was consistently reduced to less than 0·5 km sec^{-1}. It seems more reasonable to assume that the echoes are associated with some system moving permanently with the earth.

2. Echoes of long delay were observed mainly or wholly during the daytime. Galle, Talon and Ferrié [8] state that they disappeared during a solar eclipse, and reappeared slowly afterwards. This observation suggests that the formation of the echoes depends on the ultra-violet light from the sun, and can be affected by the shadows of the earth and the moon.

3. The echoes have been heard on frequencies 9·5 and 6·7 Mc/s, but not on frequencies 13·5 and 20·7 Mc/s. This suggests that some feature of the echoes may be associated with the ionosphere. Galle, Talon and Ferrié used a frequency of 12 Mc/s, but their observations were made in Indo-China, that is, near the equator, whereas all other reports of long delayed echoes came from temperate latitudes.

We suggest, very tentatively, that long delayed echoes are due to the propagation of guided waves of the type described by Pedersen [6] (see section 3), over long curved paths formed by belts of ions outside the earth but fixed relative to it, and that the successful launching of such waves depends on the state of the ionosphere.

11. Suggestions for Future Tests

In view of the conclusions of the last section, it is suggested that in any further attempts to observe echoes of long delay, it is desirable to use rather lower frequencies than were used in our tests. The frequency chosen should probably be less than 10 Mc/s, and it may also be undesirable to use aerials whose directions of maximum transmission or sensitivity are radially outwards from the earth. Better results might be achieved with aerials having horizontal directivity, and if a study could be made as to how the strength of the echoes depends on the directivity of the aerials, this would give some useful information about the nature of the waves responsible for the echoes.

Acknowledgements—The work described in this paper was carried out partly at the *Solar Physics Observatory, Cambridge*, and partly at the *Cavendish Laboratory, Cambridge*, where it formed part of a programme of radio research supported by the *Department of Scientific and Industrial Research*. Our grateful thanks are due to *Messrs. Cable and Wireless, Ltd*, and especially to their Chief Engineer, Mr. Smale, and to Mr. Keen and the staff of the transmitting station at Ongar, for arranging the special transmissions. We also wish to thank our assistant, Mr. A. B. Reeder, who gave invaluable help with the construction of the receivers, and the taking of the observations.

References

[1] Størmer, C.; Nature 1928 **122** 681. [2] Van der Pol, B.; Nature 1928 **122** 878. [3] Fuchs, J.; Hochfreq. Tech. und Elek. 1934 **44** 163. [4] Fuchs, J.; *ibid* 1935 **45** 111. Galle, J. B.; Onde Électrique 1930 **9** 257. [5] Størmer, C.; Nature 1929 **123** 16. [6] Pedersen, P. O.; Det Kongel. Danske Vidensk. Selskab., Math-fys. Medd. 1929, IX, 5. [7] World Radio, 1934 **19**, Series of articles, mostly under the heading "World Radio Research League: Secretary's Report". [8] Galle, J. B., Talon and Ferrié; Comptes Rendus 1930 **190** 48. [9] Appleton, E. V.; Nature 1928 **122** 879. [10] Breit, G.; Proc. Inst. Radio Eng. 1929 **17** 1508. [11] Thomas, L. H.; Nature 1929 **123** 166. [12] Joos, G.; Hochfr. Tech. u. Elek. 1931 **37** 136. [13] Størmer, C.; Naturwiss. 1929 **17** 643. [14] Størmer, C.; Comptes Rendus 1929 **189** 365. [15] Janco, N.; Proc. Inst. Radio Eng. 1934 **22** 923. [16] Jelstrup, H. S.; Onde Électrique 1928 **7** 538. [17] Jonescu, Th. and Mihul, C.; Comptes Rendus 1934 **199** 1389. [18] Chapman, S.; Journ. Geophys. Res. 1950 **55** 361. [19] Chapman, S. and Ferraro, V. C. A.; Terr. Mag. and Atm. Elec. 1931 **36** 77 and 171. [20] Chapman, S. and Ferraro, V. C. A.; Terr. Mag. and Atm. Elec. 1932 **37** 147 and 421. [21] Chapman, S. and Ferraro, V. C. A.; Terr. Mag. and Atm. Elec. 1933 **38** 79. [22] Dostal, H.; Ann. der Phys. 1932 **14** 971. [23] Chapman, S. and Bartels, J.; Geomagnetism. Oxford University Press, 1940.

VOL.
2
1952

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 282 to 291. Pergamon Press Ltd., London

A radio echo method for the investigation of atmospheric winds at altitudes of 80 to 100 km

J. S. Greenhow

Jodrell Bank Experimental Station, University of Manchester

(Received 14 March 1952)

ABSTRACT

The effects of high altitude winds on the ionized trails produced by meteors are discussed. The variations of wind velocity with height between altitudes of 80 to 100 km can be determined from the fluctuations in amplitude and the changes in width of the radio echoes from the trails. These variable winds are present at all times in the meteor region, and the velocities at points separated by only 5 to 10 km may differ by as much as 50 m/sec. Uniform wind velocities can also be measured by observation of the steady drifts in range of long duration echoes. These uniform winds also have velocities of the order of 50 m/sec. A reversal in the predominant wind direction is observed between 0100 and 0200 hrs local time, the direction prior to 0100 hrs being roughly from North to South, and after 0200 hrs from South to North.

I. Introduction

Visual observations of the movements of noctilucent clouds and persistent meteor trails have demonstrated that high velocity winds exist in the 100 km region above the earth's surface. The non-uniformity in these wind motions has been determined by Störmer [1] and others by visual observation of the progressive distortions which take place in long duration meteor trails. Applications of radio techniques to these problems have been made by Mitra [2], Manning, Villard and Peterson [3], Ellyett [4] and Greenhow [5], [6]. The present paper describes further experiments which were designed to investigate the effects of these upper atmosphere winds upon the trails of ionisation produced by meteors. From the amplitude fluctuations and pulse broadening observed in the radio echoes from meteor trails, it has been possible to estimate the magnitude of the non-uniform winds at altitudes of 80–90 km. A certain amount of information on uniform wind velocities has also been obtained.

II. Technique

The transmitter used in the experiments operated on a wavelength of 8·4 m, the peak power being 150 kw at a pulse repetition frequency of 150 c/s. The aerial system, which was used both for receiving and transmitting, consisted of a horizontal dipole and parasitic reflector at a height of half a wavelength above the ground. The aerial was orientated to give a wide beam in a northerly direction. Two separate equipments were used for recording the amplitude fluctuations in the meteor echoes and for observing the detailed shape of the echo pulses.

(a) For the observation of the amplitude fluctuations the receiver output was presented as a displacement on a cathode ray tube and, when a meteor echo appeared, the amplitude variations were recorded photographically by a camera through which the film was driven at a speed sufficient to resolve successive echo pulses. The first echo pulse to be received from a meteor trail was used to switch on the camera motor, which required 0·1 sec to accelerate to its operating speed. The motor was stopped manually on the disappearance of the echo.

(b) For the observation of the pulse shape and range drifts a cine camera was used to photograph a range-amplitude display. In this experiment it is necessary to resolve range differences of fifty metres, requiring a stability in range of at least ten metres, or in time of better than 0·1 μs. Since meteor echoes can be received at ranges up to 1000 km the equipment must therefore be stable to within 1 part in 100000. The various triggering pulses were all derived from a 150 K/cs crystal oscillator, spaced 150 c/s pulses being obtained by division. These pulses were used to initiate the modulator of the transmitter. The stability requirement of 0·1 μs for the radio frequency pulse corresponds to only 3·6 oscillations at a wavelength of 8·4 m. In the ordinary self excited oscillator stage of the transmitter used in previous experiments [5], [6], there was an uncertainty in the delay time between the application of the modulator pulse, and the commencement of oscillation, of approximately $\pm$ 0·5 μs. A low power continuous wave oscillator was therefore coupled to the oscillatory circuit in the transmitter in order to assist the initiation of the main oscillation. By this method the required degree of stability was achieved. The pulse generated by the transmitter was approximately 3 μs in duration.

Range differences of the order of 50 m were measured by displaying the meteor echoes on a very fast triggered time base, 5 km in length. One kilometre range markers were produced directly from the 150 Kc/s waveform, and the time base was triggered from every fifth 1 km range marker. At a pulse repetition frequency of 150 c/s successive transmitter pulses were separated by 1000 km, so that an echo pulse was present on only one in every 200 of the time base sweeps. It was therefore necessary to reduce the intensity of the horizontal sweep relative to that of the echo pulses, in order to prevent fogging of the film. This was achieved by using an automatic strobe to brighten the cathode ray tube trace over a range of a few kilometres on either side of the echo.

The cathode ray tube was photographed by a cine-camera taking one frame every 0·6 sec, with an exposure time of 0·5 sec. A number of echo pulses were therefore superimposed on each frame. These cine records were correlated with the continuous film records of the amplitude fluctuations described in (a).

III. The Effect of High Altitude Winds Upon Trails of Meteor Ionisation

(a) *Fluctuations of echo amplitude*

Observations of radio echoes from the trails of ionisation produced by meteors have shown that the majority of echoes are obtained only when the trails are at right angles to the direction of observation [7]. A large scale uniform wind motion would cause a trail to drift parallel to itself, the slant range of the radio echo changing at a rate equal to the line of sight component of the wind velocity. The amplitude of the radio echo from such a trail would decay to zero without fluctuating as the column of ionisation diffuses. Previous work has shown, however, that these ideal conditions are rarely encountered because the trails quickly become distorted due to irregular winds (Greenhow [5], [6]). Echoes may be received from several reflecting centres along the trail, producing fluctuations in the resultant echo amplitude by interference effects.

If echoes of amplitudes a_1, a_2, . . . a_n, are received from n reflecting centres, moving with velocities v_1, v_2 . . . v_n in directions θ_1, θ_2 . . . θ_n to the line of sight,

VOL. 2 1952

then the resultant echo at 0 (Fig. 1) is given by an expression of the form

$$Y = \sum_{1}^{n} a_r \sin(\omega t + \delta_r)$$

where ω is the angular frequency of the radio waves.

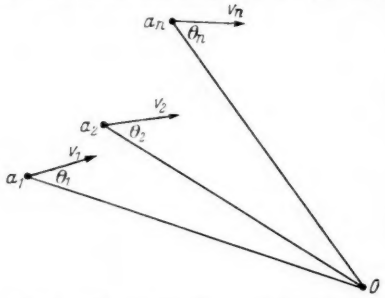


Fig. 1. The reflection of radio waves from a meteor trail consisting of n reflecting centres in relative motion.

This can be rewritten

$$Y = A \sin(\omega t + \Delta)$$

where

$$A^2 = \sum_{r=1}^{n} \sum_{s=1}^{n} a_r a_s \cos(\delta_r - \delta_s).$$

Now the change in phase δ_r of the wave reflected from rth centre in time t is

$$\delta_r = \frac{2\pi}{\lambda} 2 V_r \cos\theta_r \,.\, t$$

or twice the line of sight movement of the centre.

$$\therefore\ \delta_r - \delta_s = \frac{4\pi}{\lambda} V_{rs} \,.\, t$$

where V_{rs} is the relative line of sight velocity of any two centres

$$\therefore\ A^2 = \sum_{r=1}^{n} \sum_{s=1}^{n} a_r a_s \cos \frac{4\pi V_{rs}}{\lambda} t\,.$$

When $n = 2$

$$A = \left(a_1^2 + a_2^2 + 2 a_1 a_2 \cos \frac{4\pi V_{12}}{\lambda} t\right)^{\frac{1}{2}}.$$

The fluctuations have a regular period, $T = \dfrac{\lambda}{2V_{12}}$ between successive maxima. As n increases the fluctuations become complex in character, and it has been shown that in a long duration meteor echo a Rayleigh distribution of amplitudes is eventually formed, consistent with the theory of a large number of reflecting centres moving with random velocities (GREENHOW [6])*.

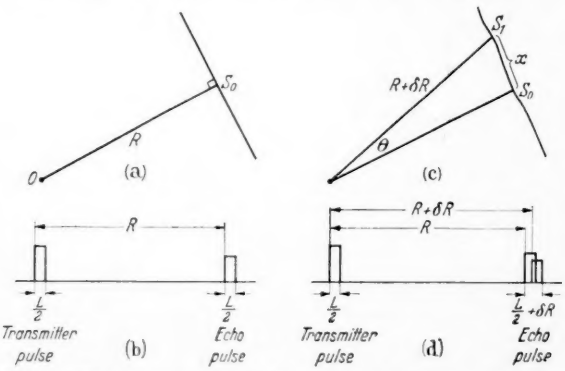


Fig. 2. The appearance of the radio echo from a meteor trail on a range-amplitude display.
(a) and (b) The echo from an initially linear trail.
(c) and (d) The echo from a trail after distortion has commenced.

Regular fluctuations of the type illustrated in Fig. 5(d) will only be observed when two reflecting centres predominate, when the period of fluctuation gives the relative line of sight velocity of the centres.

(b) *Variations in echo width*

The trail distortion discussed in the previous section will also

* It has been pointed out by Mr. J. A. RATCLIFFE that the complex fluctuations observed in very long duration meteor echoes are capable of an alternative explanation, in terms of the steady drift of an irregular diffracting screen [10]. Only meteor echoes showing regular fluctuations have been used in the following analysis, and in this case the interpretation of the fluctuations as due to a relative movement of reflecting centres appears to be unambiguous.

produce a variation in the width of the resultant echo from a column of meteor ionisation. Before distortion occurs, a discrete echo will be obtained from the section of a trail near the foot of the perpendicular at S_0 [Fig. 2(a)]. On a range amplitude display this echo would appear at a range R with a width equal to that of the transmitted pulse [Fig. 2(b)]. As the trail distorts an echo may be received from another point S_1, some distance along the trail from S_0, at a range of $R+\delta R$. On a range amplitude display the resultant echo would now be composite, having a width exceeding that of the transmitter pulse by an amount δR [Figs. 2(c) and (d)]. If the effect of variable winds is to produce irregularities along the trail, rather than to cause it to turn as a complete unit, then for $R \gg \delta R$ $x^2 = 2R\,\delta R$, where x is the separation of the reflecting centres. The increase in width of a complex echo over the width of the transmitter pulse therefore enables the separation of the reflecting centres to be determined approximately. The angular distortion of the section of trail at S_1, [Fig. 2(c)] must be approximately equal to θ, in order to satisfy the specular reflection condition, and it is given by $\theta \sim \frac{x}{R}$.

IV. Experimental Results

(a) *Increase of distortion with echo duration*

VOL. 2 1952

The experimental observation of meteor echoes on a high speed time base, using the techniques described in II(b), has shown that in many cases the width of the echo pulses are greater than those from a discrete reflecting centre. The variation of maximum echo width, and the angular distortion derived from it, are given as a function of echo duration in Fig. 3.

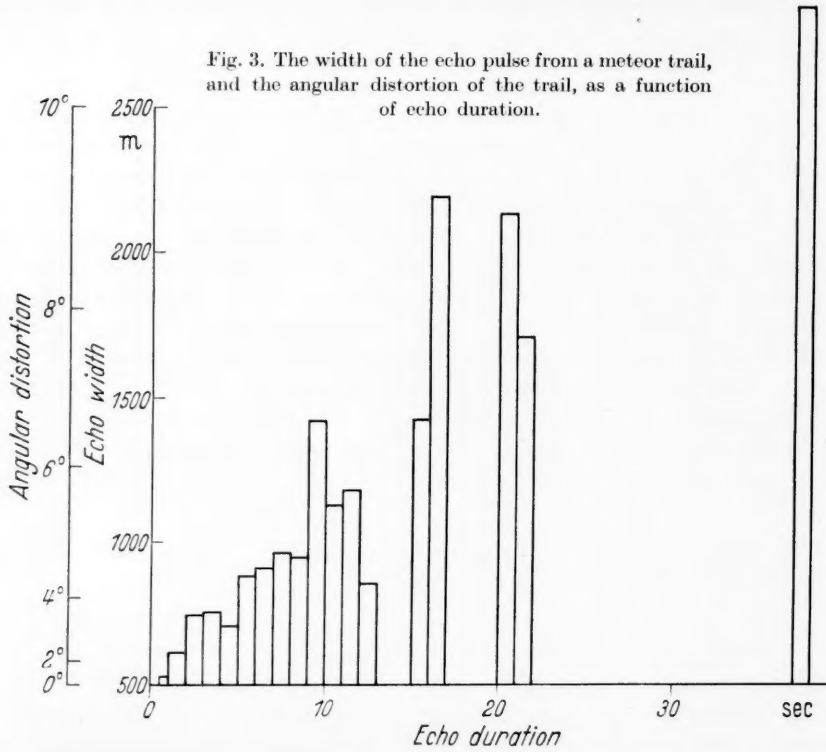
In general, echoes with durations less than 0·5 sec do not distort sufficiently to give a measurable increase in received pulse width, although fluctuations in amplitude may be observed, indicating that some distortion must occur within this time. Figs. 4(a) and (b) illustrate an echo which does not fluctuate in amplitude, its width being the same as that of an echo from a discrete reflector. For almost every echo with a duration greater than 0·5 sec, however, there is an increase in received pulse width, showing that these echoes are composed of a number of separate reflections (Figs. 5 and 6). The corresponding continuous film displays invariably show fluctuations in echo amplitude.

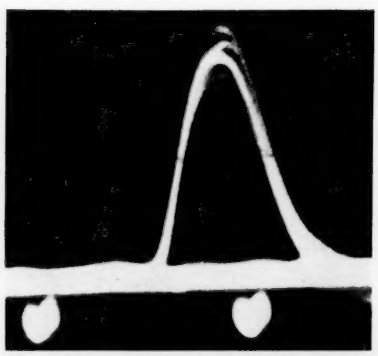
Fig. 3 shows that the amount of echo broadening increases with duration, becoming approximately 2 km for echoes lasting 30 sec. As the time after the passage of the meteor increases, sections of the trail at increasing distances from S_0 will distort sufficiently to give a detectable echo, thus accounting for the observed increase in echo width with duration.

The mean distortion of meteor trails with durations of up to 2 sec is less than 3 degrees. As this includes more than 80% of all echoes at a wavelength of 8·4 m, the assumption that radio echoes from meteor trails are obtained by specular reflection is approximately valid for the majority of echoes at this wavelength.

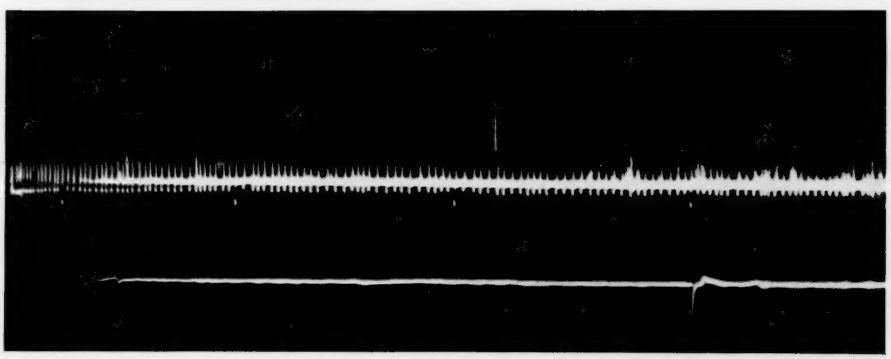
For a trail in which $\delta R > 450$ m, the individual echoes may be returned as completely separate components. Several frames of a long duration echo which persisted for 37 sec are shown in Fig. 6. On a normal range amplitude display used for meteor observations, with a time base several hundred kilometres in length and using a wide transmitter pulse, this echo would appear as two or three discrete

Fig. 3. The width of the echo pulse from a meteor trail, and the angular distortion of the trail, as a function of echo duration.





(a)



(b)

Fig. 4. (a) The appearance of a short duration meteor echo using high range resolution. The individual echo pulses are superimposed. 1 km range markers are displayed as downward deflections. (b) The continuous film record of the same echo in which successive echo pulses are resolved, showing the absence of pronounced amplitude fluctuations.

enduring components. When presented on a high speed time base, however, the echo is seen to be continually changing in character. On the frames shown, eight different components can be observed, most of which last for only a few seconds. At times the echo is spread in range by as much as 3 km, and reflections must be

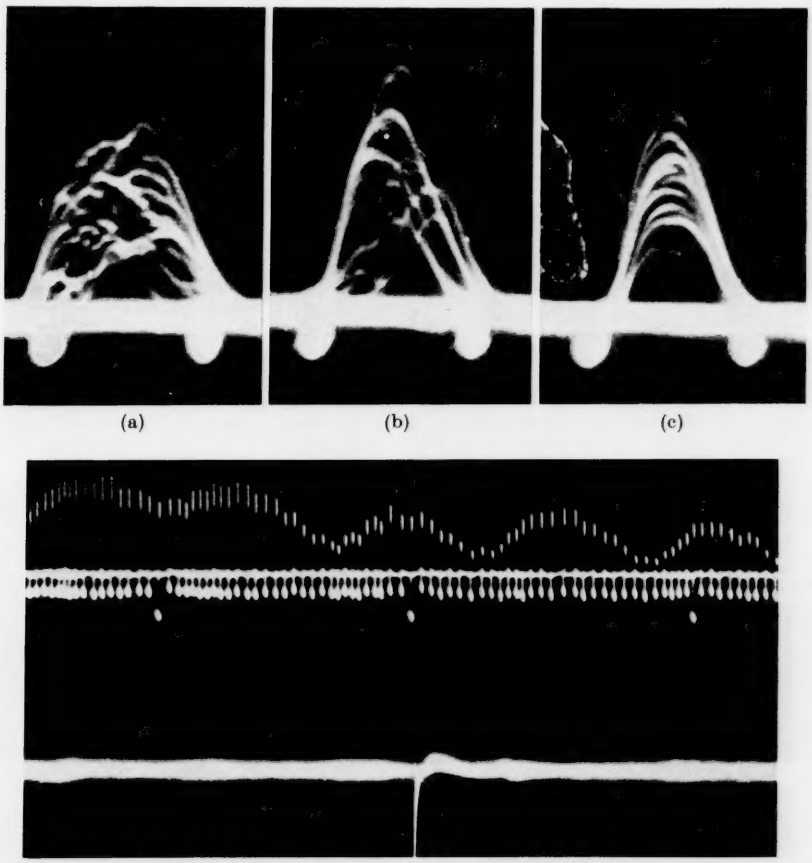


Fig. 5. (a), (b) and (c) Three stages in the life of a long duration meteor echo, showing the variation in width of the echo. (d) Part of the continuous film record of this echo in which the amplitude undergoes fluctuations with a period of approx. 0·1 sec.

taking place from a distorted length of trail of the order of 30 km in length. The components *A* and *H* are exceptional, the former was the first to appear and, during its lifetime of almost 25 sec, it drifted in range by approximately 1 km*. Component *H* which appeared 30 sec from the beginning of the echo drifted approximately 0·5 km in range during its lifetime of 7 sec.

* After 15 sec echo *A* was beginning to drift off the time base and, in order to offset this movement, the time base trigger pulses were advanced by 1 km. The whole echo pattern is therefore apparently displaced by 1 km at this time. Corresponding range calibration pulses have been marked to assist in identification.

VOL. 2 1952

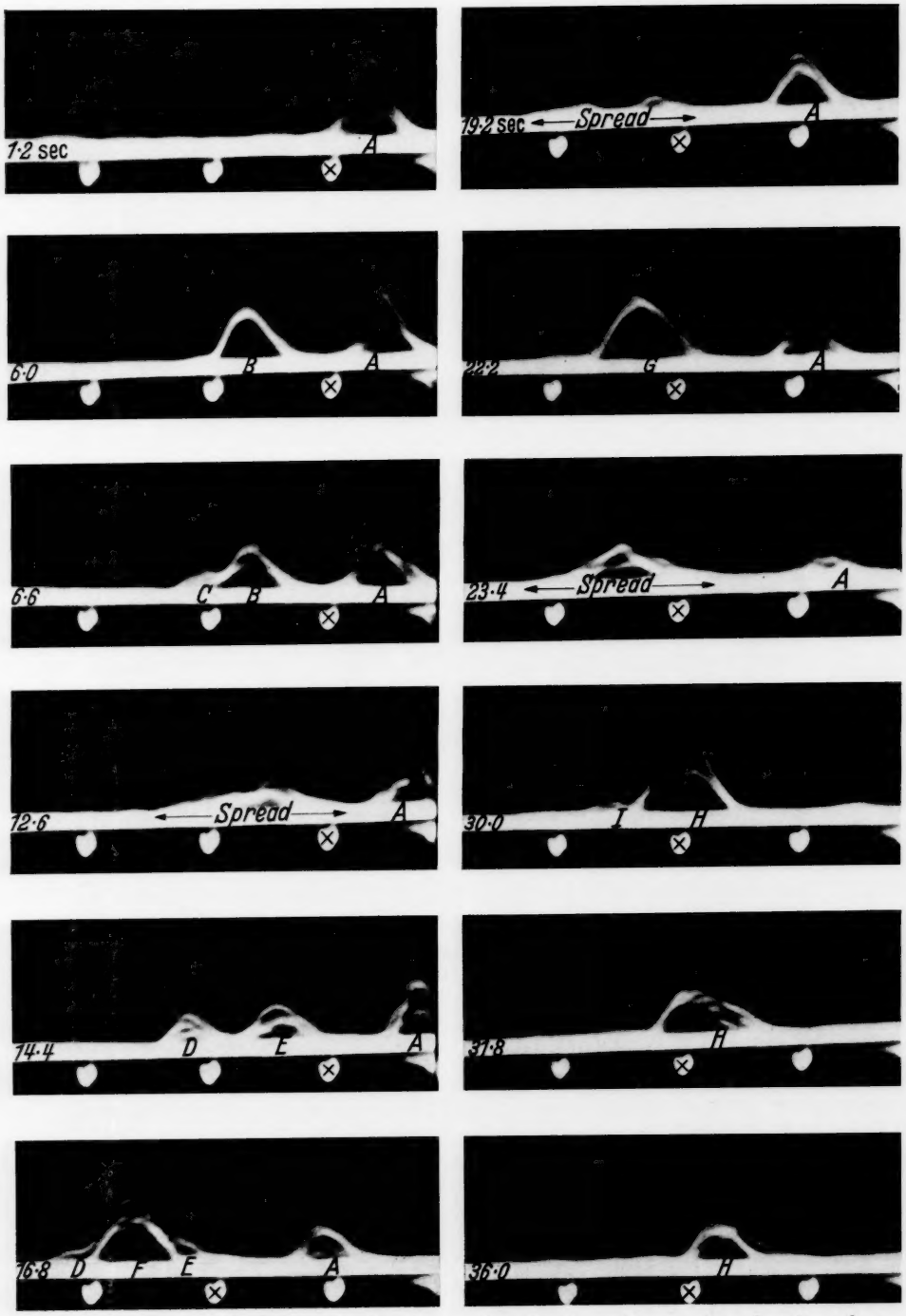


Fig. 6. A series of cine frames showing the complex structure of a very long duration echo.

(b) *Correlation between amplitude fluctuations and increase in echo width*

If the amplitude fluctuations in a meteor echo are produced by a relative movement of reflecting centres, then when a series of n regular fluctuations is observed, there should be a simultaneous change in echo width of $n\lambda/2$. As the uncertainty in determining the width of a complex echo is of the order of 50 m, owing to its variation in shape, this effect would only be observed in an echo which remained sufficiently stable to give a long series of regular fluctuations. Several frames of an echo which decreased in width by approximately 280 m in 7 sec are shown in Figs. 5(a), (b) and (c). During this period of 7 sec there were approximately 60 fairly regular

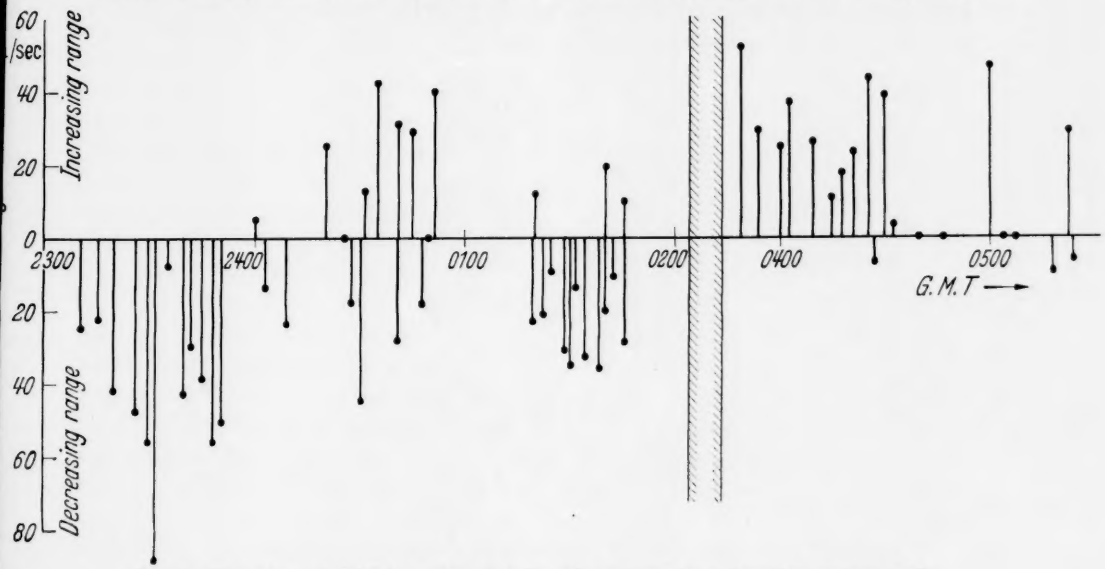


Fig. 7. The rates of drift in range of radio echoes from long duration meteor trails.

amplitude fluctuations, [Fig. 5(d)] corresponding to a change in the line of sight separation of two centres of 252 m at a wavelength of 8·4 m. This figure is in reasonable agreement with the observed change in echo width, supporting the association between the amplitude fluctuations in a meteor echo and a relative movement of reflecting centres.

(c) *Measurement of wind velocities*

(i) *Uniform wind velocities*—Observations using the high range resolution equipment were made during the nights of Dec. 12th–13th, and 13th–14th, 1950, when the Geminid meteor shower was active. Approximately 70% of the radio echoes with durations greater than 0·5 sec occurring during this shower lie between altitudes of 83 and 90 km, and the wind measurements described in this paper therefore refer to this altitude range. 62 individual meteor echoes, or components of complex echoes, giving measurable range drifts were obtained. The drifts were only accepted if a steady motion of the echo was observed over several cine frames, so as to exclude spurious range drifts due to a change in the reflecting point along the trail (Ellyett [4]). The rates and sense of drift for these 62 echoes are shown in Fig. 7, together with the times of the measurements.

VOL. 2 1952

Between 2300 and 2400 hrs G.M.T. only decreasing range drifts were observed and, as echoes are received from a northerly direction with the aerial system used, these results indicate a wind with a component of 30–60 m/sec blowing from a northerly to a southerly direction. Between 2400 and 0200 hrs both increasing and decreasing range drifts were recorded. This result could be due either to an east-west (or west-east) wind blowing across the wide aerial beam, or in the absence of a substantial uniform drift, to turbulence. No observations were made between 0200 and 0320 hrs, but after 0320 hrs the majority of the echoes increased in range with drift velocities between 40 and 60 m/sec. The wind therefore appears to have reversed its direction prior to 2400 hrs and it then had a large south to north component. As measurements have been made on only two nights, more results will be required in order to decide whether or not the observed wind reversal is general throughout the year.

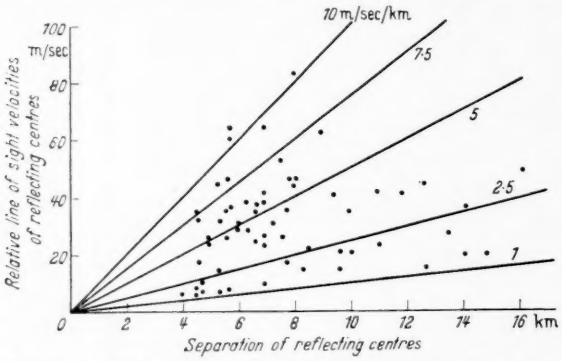


Fig. 8. Wind gradients in the upper atmosphere. Any point in the figure represents the difference in wind velocities at two points separated by a given distance.

The wind measurements after 0300 hrs are similar in magnitude and direction to the measurements made by MANNING, VILLARD and PETERSON [3] who found that the predominant wind direction was from SSW to NNW; the mean velocity being 30–40 m/sec. No results by these workers are available for comparison before 0300 hrs. MITRA [2] has observed a reversal in wind direction between 2400 and 0100 hrs from ionospheric measurements, but in this case the wind directions are directly opposite to the ones described in this paper, the reversal taking place from approximately south to north, to north to south. More recent measurements by PHILLIPS [8], and SALZBERG and GREENSTONE [9] using MITRA'S technique, also indicate winds with directions opposite to those obtained from meteor observations. The exact heights of these ionospheric measurements are uncertain, and whereas the winds deduced from the behaviour of meteor trails definitely refer to altitudes of 80–100 km, some of the ionospheric results may apply to the *F*-region at appreciably greater heights.

(ii) *Non-uniform winds*—The complex phenomena observed in radio echoes from the columns of ionisation produced by meteors show that the primary effect of upper atmosphere winds upon the form of an echo is due to trail distortion, rather than to uniform translation. It has been shown in III(a) that measurements of the degree of non-uniformity of the winds can be made from meteor echoes which fluctuate in amplitude with a regular period, that is those for which only two main reflecting centres are present. The period of fluctuation then gives the relative line of sight velocity of the reflecting centres, and the separation of the centres can be found from the increase in echo width over the width of an echo from a discrete reflector. Each individual determination of the velocity difference of two centres is plotted against their measured separation in Fig. 8. The sloping lines in Fig. 8 correspond

to various values of the wind gradient. Most of the wind gradients lie between 1 and 7·5 m/sec/km, although the maximum value observed is greater than 10 m/sec/km. Fig. 8 illustrates the extent of the non-uniformity of the winds at heights of 80–90 km. The wind velocities at points separated by only 5 to 10 km can differ by more than 50 m/sec, which is of the same order as the uniform wind velocities discussed in the previous section. These values are similar in magnitude to those obtained from observations of the distortions occurring in visual meteor trails. STÖRMER [1] has observed variations in wind velocity of some tens of metres/sec over height intervals of the order of 10 km.

The effects of high altitude winds upon the columns of ionisation produced by meteors are illustrated by the complex echo in Fig. 6. The presence of a uniform wind drift is shown by the steady change in range of the echo components *A* and *H*, while the effect of variable winds is to cause echoes from other sections of the trail to rapidly change in character, individual components remaining stable for only a few seconds. The rapid changes in the speeds and directions of the uniform winds measured by MANNING, VILLARD and PETERSON are probably due to this large variation of wind velocity with height, as these workers measured comparatively small range drifts in localised regions of the atmosphere.

Acknowledgements—The work described in this paper was carried out at the *Jodrell Bank Experimental Station* of the *University of Manchester*. The author wishes to thank Professor A. C. B. LOVELL, the Director of the Station, for his constant interest and advice, and also Mr. J. G. DAVIES and Miss M. ALMOND for their assistance in taking the experimental observations. He is also indebted to Mr. J. A. RATCLIFFE of the Cavendish Laboratory for a most helpful discussion on the problem of the diffraction of radio waves at an irregular screen. The work was made possible through financial assistance from the Department of Scientific and Industrial Research.

VOL. 2 1952

REFERENCES

[1] STÖRMER, C.; Astrophysica Norvegica 1937 **3** No. 5. [2] MITRA, S. N.; Proc. Inst. Elect. Engrs 1949 (Pt. III) **96** 441. [3] MANNING, L. A., VILLARD, O. G., PETERSON, A. M.; Proc. Inst. Radio Engrs 1950 **38** 887. [4] ELLYETT, C. D.; Phil. Mag. 1950 **41** 694. [5] GREENHOW, J. S.; Phil. Mag. 1950 **41** 682. [6] GREENHOW, J. S.; Proc. Phys. Soc. B 1952 **65** 169. [7] LOVELL, A. C. B., BANWELL, C. J., CLEGG, J. A.; Mon. Not. Roy. Ast. Soc. 1947 **107** 164. [8] PHILIPS, G. J.; Observatory 1951 **71** 105. [9] SALZBERG, C. D., GREENSTONE, R.; J. Geophys. Res. 1951 **56** 521. [10] BOOKER, H. G., RATCLIFFE, J. A. and SHINN, D. H.; Phil. Trans. Roy. Soc. A 1950 **242** 579.

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 292 to 300. Pergamon Press Ltd., London

The relation between point discharge current and field

J. Alan Chalmers, M.A., Ph.D.

Senior Lecturer in Physics, Durham Colleges in the University of Durham

(Received 5 April 1952)

ABSTRACT

The empirical relation $I = A(F^2 - M^2)$ of Whipple and Scrase [1] is derived theoretically on certain assumptions involving space charges. The factors concerned in the values of A are discussed.

1. Introduction

Whipple and Scrase [1], investigating the currents produced by point discharge during periods of high electric field in the atmosphere, found an empirical relation between the point discharge current, I, through a single point and the field, F, at the earth's surface nearby:

$$I = A(F^2 - M^2) \tag{1}$$

where A and M are constants.

Other workers have found a similar relation and Hutchinson [2] has accounted for apparent deviations from the formula.

Peek [3], who measured corona discharge from a high-tension line, found the effect to be proportional to $(e - e_0)^2$, where e is the voltage on the line. While not quite similar to measurements of point discharge in a naturally occuring field, these results are clearly related.

Hitherto it seems that no attempt has been made to derive eq. (1); the present work derives (1) under certain assumptions and conditions.

2. Assumptions

In the first place, it is assumed that electrical conditions in the atmosphere are stationary, that is to say that the electric fields at all points are not altering with time. Whipple and Scrase [1] selected periods when conditions were as nearly as possible stationary.

It is assumed that the single point, through which the current is measured, can be considered as one of a large number of similar points all of the same height h and arranged in rectangular array with spacing d, the whole extending over a wide area.

It is assumed that the cloud giving rise to the field which causes point discharge is widespread, so that the lines of force are vertical and the equipotential surfaces horizontal except when close to the points.

It is assumed that the production of ionisation by collision gives rise to small ions of both signs and that combination of these ions to form large ions is negligible during the times taken for them to reach the cloud.

For convenience in discussion, we consider the case where the field is positive, so that there are negative ions moving upwards from the points.

It is assumed that the place where F is measured is fairly representative of the area between the points.

There is, as will be seen, no need to make any assumption as to the processes at work in the production of ionisation by collision. It is assumed that the current so produced is limited by space charges in much the same way as the thermionic current in a valve is limited by space charge; in fact, the present deduction bears many points of resemblance to LANGMIURs deduction of the 3/2 power law, the difference lying in the fact that electrons in a vacuum are continuously accelerated in a field, while ions in air have a definite steady velocity in a field.

It is assumed that the wind can be neglected.

3. SPACE CHARGE THEORY

Though the lines of force are parallel when well above the points, they tend to concentrate on the points and give, in the immediate vicinity of each point, a volume in which the field is so strong that ionisation by collision can occur. The positive ions so produced will travel into the point and provide the measured point discharge current; the corresponding negative ions will travel upwards along the lines of force (in the absence of wind) and will form a space charge in the air above the points. Some of the lines of force will end on these negative ions and so the number reaching each point is reduced, limiting the volume in which ionisation by collision can occur. Clearly a balance will be set up so that the volume in which ionisation by collision occurs is limited by the space charge to give just the current to provide this space charge; then we need not discuss further the process of ionisation by collision or even the size of this volume, provided it is small.

VOL. 2 1952

Let us consider a volume bounded at its ends by two equipotential surfaces, and at its sides by lines of force. Let this volume contain negative ions of charge $-e$ and mobility w. In the absence of wind, the current will flow along the lines of force and so, in the stationary state, any charge entering at one end of the volume will leave it at the other end and the same current flows through each cross-sectional area.

If we assume the field X to be uniform across an equipotential cross-section, and if n is the number of ions per unit volume, then for an area S, the current I is

$$I = n\,e\,w\,X\,S \tag{2}$$

towards the negative pole of the field (actually a negative current moves in the opposite direction).

Applying GAUSS' theorem to the volume between two cross-sections distant dx apart:

$$\frac{d}{dx}(X\,S) = 4\pi\,n\,e\,S. \tag{3}$$

Combining (2) and (3)

$$\frac{4\pi I}{w} = X\frac{d}{dx}(X\,S).$$

Multiplying by S and integrating between areas S_1 and S_2, where the fields are X_1 and X_2 respectively

$$(X_1\,S_1)^2 - (X_2\,S_2)^2 = \frac{8\pi I\,v}{w} \tag{4}$$

where v is the volume between S_1 and S_2.

Where the field is not uniform over the cross-sections of the volume concerned, (4) can be applied to volumes of small cross-section and the results added, giving:

$$\sum (X_1 S_1)^2 - \sum (X_2 S_2)^2 = \frac{8\pi}{w} \sum I v .$$

(4) is the general formula relating point discharge current and field in the absence of wind and is believed to be a new result. If we now consider plane surfaces, not necessarily equipotentials, for S_1 and S_2, and if X is now the component of field perpendicular to the surface, we get the same eq. (4).

In the special case where the volume concerned is a cylinder of height H and cross-section S, (4) reduces to

$$X_1^2 - X_0^2 = \frac{8\pi}{w} \frac{I}{S} H \tag{5}$$

which was deduced by WILSON [4] and applied to conditions in the atmosphere well above the points. If H is the height above the earth, then X_0 is the field which would exist at the earth if the lines of force remained parallel right down to the earth, with the same current density. Clearly X_0 is not the same as F, the field that actually occurs at the earth but is related to it. (5) has been rewritten by WHIPPLE and SCRASE [1] as:

$$X_1^2 = \frac{8\pi}{w} \frac{I}{S} (H + c). \tag{6}$$

Where $-c$ is the height at which X would be zero if the conditions of parallel lines of force continued down. It is clear that both X_0 and c may depend upon I.

4. POINTS IN RECTANGULAR ARRAY

With a number of similar points in rectangular array, the space charge from each point will spread out into what may be called its "space charge volume" as the ions move upwards, and at a sufficient height the space charge streams from all the points will have merged to give a uniform arrangement of ions with vertical lines of force and horizontal equipotentials, the conditions under which (5) holds. If we follow these lines of force downwards, some will end on the upward moving negative ions, some on the points and some on the ground between the points; it is these last which give rise to the measured field at the earth's surface, and if we can determine, at some height, what proportion of the lines of force go to the earth, we can find the relation between point discharge current and field at the earth's surface.

It does not, however, seem possible to determine the true shape of the space charge volumes, and all we can do is to make approximate assumptions. In actual fact, at a sufficient height above the points, potential, field and space charge are uniform over a horizontal plane. The assumption we shall make, to give a basis for calculation, is that this is true at heights above $(h + a)$, where h is the height of each point, and that below this level the fields and potentials both within and outside the space charge volumes vary in such a way as to give uniformity of field and potential at $(h + a)$; since space charge exists only within and not outside the space charge volumes, we cannot get uniformity of space charge at $(h + a)$. Actually, wind and diffusion will gradually produce uniformity of space charge, and our

assumption is that below $(h + a)$ the space charge is wholly within the space charge volumes and above $(h + a)$ it is evenly distributed.

There are two ways in which it is possible to make assumptions in regard to the space charge volumes. In the first, it is assumed that the space charge volume of a point is exactly the same shape under all conditions, *i.e.* that the different space charge volumes do not affect one another. Or we can assume that the space charge volumes affect one another so that they touch at the height $(h + a)$ at which fields and potentials are uniform.

5. Conical Space Charge Volumes

First, it is assumed that the space charge volume of every point is a cone of semi-vertical angle θ, the apex of the cone being at the point. Since the ions at the outside of the space charge volume follow lines of force, the surface of the cone must contain lines of force. It is also assumed that the vertical component of electric field is uniform across the cone. Further, it is assumed that the volume in which ionisation by collision occurs is very small compared with the volume of the cone up to height $(h + a)$.

The conditions are suitable for the application of (4); at a height x above the point, $S = \pi x^2 \tan^2\theta$ and $v + v_0 = \frac{1}{3} S x$ where v_0 is the volume in which ionisation by collision occurs.

VOL. 2 1952

Then,

$$(X \pi x^2 \tan^2\theta)^2 - (X_0 S_0)^2 = \frac{8\pi I}{w}\left(\frac{\pi}{3} x^3 \tan^2\theta - v_0\right) \tag{7}$$

where X_0 is the vertical componant of the field over the area S_0 which is the upper boundary of the volume v_0.

For the time being we can neglect the small quantities $(X_0 S_0)^2$ and $\frac{8\pi I}{w} v_0$ and get:

$$X^2 = \frac{8I}{3w \tan^2\theta} \frac{1}{x}. \tag{8}$$

At the height $(h + a)$

$$X_a^2 = \frac{8I}{3w \tan^2\theta\, a}. \tag{9}$$

Integrating (8) from $x = 0$ to $x = a$,

$$V_a^2 = \frac{8I}{3w \tan^2\theta} 4a. \tag{10}$$

$$\therefore \frac{V_a}{X_a} = 2a. \tag{11}$$

Now, considering the volumes outside the space charge volumes, since there is no space charge in these:

$$F d^2 = X(d^2 - \pi x^2 \tan^2\theta) \tag{12}$$

assuming F, the field at the earth's surface to be uniformly distributed over the area.

Then

$$X_a = \frac{F d^2}{d^2 - \pi a^2 \tan^2\theta} \tag{13}$$

and

$$V_a = Fh + \int_0^a \frac{F d^2\, dx}{d^2 - \pi x^2 \tan^2\theta}.$$

Putting

$$b^2 = \frac{d^2}{\pi \tan^2\theta}$$

$$V_a = F\left(h + \frac{1}{2}\, b \log_e \frac{b+a}{b-a}\right). \tag{14}$$

Then

$$\frac{V_a}{X_a} = \left(h + \frac{1}{2}\, b \log_e \frac{b+a}{b-a}\right)\left(1 - \frac{a^2}{b^2}\right). \tag{15}$$

The comparison of (11) and (15) gives a in terms of h and b. This cannot be completely worked out, but it is necessarily true that a is smaller than b, for if $a \tan\theta > \frac{d}{2}$, the space charge volumes would overlap before the height $(h + a)$ was reached. This condition is that $a < \frac{b\sqrt{\pi}}{2}$. When a/b is small, approximation from

$$\frac{2a}{b} = \left(\frac{h}{b} + \frac{1}{2} \log_e \frac{1 + \frac{a}{b}}{1 - \frac{a}{b}}\right)\left(1 - \frac{a^2}{b^2}\right)$$

gives:

$$\frac{2a}{b}\left(1 + \frac{a^2}{b^2}\right) = \frac{h}{b} + \frac{2a}{b} + \frac{2a^3}{3b^3}$$

or

$$a = h. \tag{16}$$

Comparing (9) with (13),

$$I = \frac{3w \tan^2\theta}{8} F^2 \frac{a}{\left(1 - \frac{a^2}{b^2}\right)^2}. \tag{17}$$

If, instead of (9), we use (7) with $x = a$, we get

$$\left(\frac{F\pi a^2 \tan^2\theta}{1 - \frac{a^2}{b^2}}\right)^2 - (X_0 S_0)^2 = \frac{8\pi I}{w}\left(\frac{\pi}{3} a^3 \tan^2\theta - v_0\right).$$

Now, if we assume that ionisation by collision occurs inside a small cone of height x_0, $v_0 = \frac{\pi}{3} x_0^3 \tan^2\theta$ and $S_0 = \pi x_0^2 \tan^2\theta$ and we get:

$$I = \frac{3w \tan^2\theta\, a}{8\left(1 - \frac{a^2}{b^2}\right)^2\left(1 - \frac{x_0^3}{a^3}\right)}\left(F^2 - \frac{X_0^2 x_0^4\left(1 - \frac{a^2}{b^2}\right)^2}{a^4}\right) \tag{18}$$

which is exactly of the form of (1).

The constant A is

$$\frac{3w \tan^2\theta\, a}{8\left(1 - \frac{a^2}{b^2}\right)^2\left(1 - \frac{x_0^3}{a^3}\right)}. \tag{19}$$

When a is small compared with b, and x_0^3 small compared with a^3, we have, using (16)

$$A = \frac{3w}{8} \tan^2\theta\, h. \tag{20}$$

If some other shape than the conical were chosen for the space charge volume, we should get a result exactly similar to (19) in form, but differing in detail. For example, if the space charge volumes are paraboloids of revolution for which $S = \pi k x$, then we get

$$I = A(F^2 - M^2)$$

with $A \propto k$ and independent of h or d.

6. Space Charge Volumes Touching

If we assume that the space charge volumes adjust themselves so that they just touch at the height $(h + a)$, we can calculate the relation between I and F.

For example, if the space charge volumes are conical in shape, but θ is now no longer independent of h and d, $\tan\theta = \frac{d}{2a}$ for the cones to touch at $(h + a)$ and (9) becomes:

$$X_a^2 = \frac{32\,I\,a}{3\,w\,d^2} \tag{21}$$

and (13) becomes

$$X_a^2 = \frac{F}{1 - \frac{\pi}{4}} \tag{22}$$

$b^2 = \frac{4a^2}{\pi}$ so (15) gives:

$$\frac{V_a}{X_a} = \left(h + \frac{2a}{\sqrt{\pi}} \log_e \frac{\frac{2}{\sqrt{\pi}} + 1}{\frac{2}{\sqrt{\pi}} - 1}\right)\left(1 - \frac{\pi}{4}\right). \tag{23}$$

(11) still gives $\frac{V_a}{X_a} = 2a$, so we get:

$$a = 0{\cdot}162\,h$$

and

$$I = \frac{3\,w\,d^2\,(F^2 - M^2)}{0{\cdot}239\,h}. \tag{24}$$

If, instead of the conical shape, we assume that the space charge volumes are spheroids of semi-axes $d/2$, $d/2$ and a, so that the lines of force are vertical at the height $(h + a)$, we get:

$$I = \frac{3\,w\,d^2}{3{\cdot}25\,h}(F^2 - M^2) \tag{25}$$

which differs from (24) only in a numerical factor.

Whatever assumption we may make in regard to the shapes of the space charge volumes, we shall always reach results of the same form as (9) and (13), and so by combining them we shall always get a result of the form

$$I = A(F^2 - M^2)$$

as found empirically, though the actual value of A deduced will depend upon the assumption made in regard to the space charge volume.

7. Variation of A and M

The actual variation of A with h and d can only be obtained with a definite assumption as to the shape and conditions of the space charge volumes, and, as the results differ for different assumptions, no conclusion can be drawn. But the results of the

VOL. 2 1952

last two sections always yield A proportional to w, the mobility of the ions concerned, so that we should expect A to be greater for positive fields than for negative, since the negative ions have greater mobility. Actually, WHIPPLE and SCRASE [1] found A greater in negative fields and HUTCHINSON [2] in positive.

With A proportional to w, dimensional considerations show that A/w is of the dimensions of length, as a in (19), h in (20) and d^2/h in (24) and (25). But quantities of dimensions of length include not only h and d, but also constants occuring in expressions for the shapes of the space charge volumes, such as k at the end of section 5.

Considering, on general grounds, the variation of A with h, d remaining constant, we see that as the height of the points is increased, we should expect more lines of force to reach the points and fewer to penetrate to the ground. Thus I will be increased and F decreased; thus A must increase as h increases, though we cannot obtain any power law for this variation.

To discuss variation of A with d, h remaining constant, we shall consider the same total point discharge current density; then, as we increase d, we increase the current from each point in proportion to d^2; but, at the same time, we tend to increase the number of lines of force which can penetrate to the earth and so to increase F. It seems probable that this increase in F will not be enough to counterbalance the increase in I, so that A will increase as d increases, though not as rapidly as d^2.

By similar arguments, it can be suggested that M, the minimum value of F for any discharge to occur, will decrease as either h or d increases.

In connection with the values of A and M, we can discuss an observation by CHIPLONKAR [5] who found that a system of 4 points close together gave, for the same value of F, less total current than a single point; it does not appear that this rather curious result has yet received explanation. CHIPLONKAR's system of 4 points cannot be considered quite in the same light as a decrease of d^2 by a factor of 4, since clearly the additional points are unlikely to affect the value of F, measured a little way off; so the effect on A is probably a factor $\frac{1}{4}$, and we might expect the total of the 4 points to equal that of the single point. But the decrease in d will increase M for each point and so give a less total current for the same value of F, than for a single point. CHIPLONKAR's results are not sufficiently detailed to allow verification of the idea that a change of M is responsible.

Observational results for point discharge are so scanty that it is not possible to use them with any degree of certainty to determine how A varies, especially as the values of d can often only be estimated or obtained indirectly. The only results which can be used are those of SCHONLAND [6] at Somerset West in South Africa, of WHIPPLE and SCRASE [1] at Kew and of HUTCHINSON [2] at Durham. The values of A can be obtained from the measurements, though SCHONLAND's values of I and F do not agree very well with (1). h is measured in each case, but d is known directly only in SCHONLAND's measurements; CHALMERS [7] has obtained values of d indirectly from results for rain charges both at Kew and Durham. The figures obtainable are given in Table 1, using average values for positive and negative currents for A.

Assuming the results in Table 1, and assuming that $A \propto h^p d^q$ the figures give:

$$p = 1{\cdot}7, \quad q = 1{\cdot}0\,.$$

Table 1

Author	*Place*	$A\left(\frac{\mu a}{(V/cm)^2}\right)$	$h(m)$	$d(m)$
SCHONLAND	Somerset West	9×10^{-5}	4	5
WHIPPLE and SCRASE	Kew	9×10^{-4}	8·4	14·5
HUTCHINSON	Durham	7×10^{-4}	12	6·1

While the data are too meagre to justify much reliance being placed on these results, we may say that they do agree with the kind of variation of A obtained above from general principles.

It is clearly very desirable to have more measurements of I and F, particularly where d can be directly estimated accurately. A tree in a Forestry Commission plantation would give the ideal conditions for experiment.

8. VALUES OF X_0 AND c

Using (5) and (6) at height $(h + a)$ we get:

$$X_a^2 - X_0^2 = \frac{8\pi I}{w d^2}(h + a) \tag{26}$$

and

$$X_a^2 = \frac{8\pi I}{w d^2}(h + a + c) \tag{27}$$

where X_0 is the field that would exist at the ground if conditions above $(h - a)$ extended right down, and $-c$ is the height at which the field would be zero.

Using the assumption of space charge volumes as cones of angle θ, we can substitute (9) in (26) obtaining

$$X_0^2 = \frac{8I}{3w\tan^2\theta\, a} - \frac{8\pi I}{w d^2}(h + a)$$

$$= \frac{8I}{w\tan^2\theta}\left(\frac{1}{3a} - \frac{\pi(h + a)}{b^2}\right)$$

and using (17)

$$X_0^2 = \frac{3F^2 a}{\left(1 - \frac{a^2}{b^2}\right)^2}\left(\frac{1}{3a} - \frac{\pi(h + a)}{b^2}\right) \tag{28}$$

which, if a/b is small, leads to:

$$X_0^2 \sim F^2.$$

Or, substituting (9) in (27)

$$\left.\begin{aligned} \frac{\pi}{d^2}(h + a + c) &= \frac{1}{3a\tan^2\theta} \\ c &= \frac{b^2}{3a} - h - a \end{aligned}\right\} \tag{29}$$

or

giving approximately $c = \frac{b^2}{3h} = \frac{d^2}{3\pi h\tan^2\theta}$.

VOL. 2 1952

If, instead, we consider conical space charge volumes touching, we put (21) in (26) giving:

$$X_0^2 = \frac{32\,I\,a}{3\,w\,d^2} - \frac{8\,\pi\,I}{w\,d^2}\,(h+a)$$

$$= \frac{8\,I}{w\,d^2}\left(\frac{4\,a}{3} - \pi\,(h+a)\right)$$

and using (24)

$$X_0^2 = -340\,F^2 \tag{30}$$

or putting (21) in (27):

$$\frac{4\,a}{3} = \pi(h+a+c)$$

giving

$$c = -1{\cdot}07\,h\,. \tag{31}$$

We may notice that both assumptions give c independent of I and X_0 proportional to F and thus to $I^{\frac{1}{2}}$.

9. Effect of Wind

In the above discussion, wind has been neglected, clearly the effect of wind will be to move the ions away from the assumed space charge volumes, and so allow more lines of force to reach the point, and fewer to reach the ground. Thus wind will increase I and decrease F, so increasing A. The effect will not be very large, since ions from one space charge volumes may easily be blown into another.

10. Conclusion

The assumption that space charge limits the point discharge current leads to the result:

$$I = A\,(F^2 - M^2)$$

with different assumptions as to the volumes occupied by the space charge.

The special assumptions lead to different values of A and the deduction of the variation of A from the meagre observational data cannot be very reliable.

Further measurements of I and F, in cases where the conditions approach a rectangular array of points are very desirable.

Acknowledgement—The author wishes to express his thanks to the Council of the Durham Colleges for the grant of a sabbatical year, which has allowed the present work to be pursued without interruption from other duties.

References

[1] WHIPPLE, F. J. W. and SCRASE, F. J.; Geophys Mem. 68, 1936. [2] HUTCHINSON, W. C. A.; Q. J. Roy. Met. Soc. 1951 **77** 627. [3] PEEK, F. W.; Dielectric Phenoma in High-Voltage Engineering, New York 1929. [4] WILSON, C. T. R.; Proc. Phy. Soc. 1925 **37** 32D. [5] CHIPLONKAR, M. W.; Proc. Ind. Acad. Sci. 1940 **12** 50. [6] SCHONLAND, B. F. J.; Proc. Roy. Soc. A 1928 **118** 252. [7] CHALMERS, J. A.; Q. J. Roy. Met. Soc. 1951 **77** 249.

Journal of Atmospheric and Terrestrial Physics, 1951, Vol. 2, pp. 301 to 305. Pergamon Press Ltd., London

Point discharge currents

J. Alan Chalmers, M.A., Ph.D.

Senior Lecturer in Physics, Durham Colleges, University of Durham

(Received 5 April 1952)

ABSTRACT

It is shown theoretically that the total point discharge current below a cloud is nearly independent of the nature of the points, so that point discharge currents significant in the maintenance of the earth's charge are not confined to forest and mountain areas. It is also shown that a lightning conductor will have practically no influence upon the probability of a lightning flash occurring.

VOL. 2 1952

1. Introduction

Since the pioneer work of Wormell [1] and Schonland [2], it has been realised that the process of point discharge plays an important part, perhaps the most important part, in the transfer of negative charge to the earth during storms, to balance the arrival of positive charge during fine weather.

In their correlation of the fine weather field in undisturbed conditions with the number of thunderstorms in action, Whipple and Scrase [3] included ocean as well as land storms, thus making the assumption that all storms are equally effective in transferring negative charge to the earth.

But as Whipple himself [4] remarked, one would expect very little possibility of point discharge over oceans, since the sharp points do not exist. Wichmann [5] has criticised Whipple and Scrase's inclusion of ocean storms for this reason.

The purpose of the present work is to show on theoretical grounds that the total amount of point discharge current from a given storm will be, surprisingly, nearly independent of the nature of the surface and points underneath.

2. Results for Point Discharge

Whipple and Scrase [3] found that the point discharge current, I, from a single point depends on the field, F, at the ground nearby, according to:

$$I = A(F^2 - M^2) \tag{1}$$

where A and M are constants.

The author (Chalmers [6]) has derived (1) theoretically for a rectangular array of points of height h and separation d, and has given reasons for believing that A increases and M decreases as h and d are increased.

When lines of force are vertical and equipotentials horizontal, that is to say not too close either to clouds or points, Wilson [7] showed that:

$$X_H^2 - X_0^2 = \frac{8\pi i}{w} H \tag{2}$$

where X_H is the field at height H, i the point discharge current density, w the mobility of the ions concerned and X_0 the value X would have at $H = 0$.

WHIPPLE and SCRASE [3] rewrote (2) as:

$$X_H^2 = \frac{8\pi i}{w}(H + c) \tag{3}$$

and integrated to:

$$V_H = \left(\frac{32\pi i}{9w}\right)^{\frac{1}{2}}\left((H + c)^{\frac{3}{2}} - c^{\frac{3}{2}}\right) \tag{4}$$

where c is a constant such that $X = 0$ for $H = -c$. V_H is the potential at height H.

If H represents the height of the base of a cloud, then c is small compared with H under all ordinary conditions. The author (CHALMERS [6]) has calculated c under certain assumptions and has obtained, under two different conditions either

$$c = \frac{d^2}{3\pi h \tan^2\theta}$$

where d is the separation of the points, h their height and θ the angle of the cone in which the space charge is assumed to spread, or

$$c = -1{\cdot}07\, h.$$

In the second case c is never large, while to get c comparable with H from the first case would require d^2/h to be of the same order as H; with H about 1 km and d 20 m, h would have to be of the order of centimetres or less. Hence c is always small compared with H.

For a rectangular array of points of separation d, $i = I/d^2$.

The results (2), (3) and (4) are obtained on the assumption of steady conditions and the conclusions to be drawn will be valid only on similar assumptions, but it seems that any effects of divergence from steady conditions will average out and the conclusions can be extended.

3. EFFECT OF SPACE CHARGES

If we consider variation in height of discharging points, without, at the moment, considering variation in separation, then the basis of the usual argument may be summarised as follows: the charges in the cloud give rise to a field F at the earth's surface; when F is less than the constant M in (1), there can be no point discharge. But M is increased by diminishing the heights of the points and A is decreased, so that the current per point is much decreased as the points are lowered.

This argument is, however, fallacious, since it assumes a constant F. But F is determined, not only by the conditions in the cloud, but also by the space-charges below the cloud. It is these space-charges which are concerned in the formulae (2), (3) and (4) and in the author's deduction of (1), and they cannot be left out of consideration.

The fundamental phenomenon is the separation of charge in the cloud, and it is therefore necessary to start with this. Let us assume that there is a definite rate of separation of charge in the cloud and that steady conditions are reached when the dissipation of charge just balances the separation; the dissipation of charge may occur within the cloud, affecting upper and lower charges equally, or outside the cloud, charge arriving at the lower pole being neutralised by point discharge from the ground, and that at the upper pole by a current between it and the ionosphere; if the current-voltage relations for the upper and lower dissipating currents are not

the same, then the upper and lower charges will differ in magnitude when they have attained the steady state. It is convenient to discuss the two limiting cases, when the dissipation of the lower charge occurs almost entirely inside the cloud and when it occurs almost entirely below the cloud. The actual state of affairs must lie somewhere between these two extremes.

When dissipation of the lower charge occurs almost entirely within the cloud, then the opposing processes of a constant rate of separation of charge and of a dissipation of charge proportional to the potential difference will balance for a definite potential difference between the poles of the cloud; let this be v. Below the cloud, there must be a current, i, small compared with that within the cloud, and related to the potential V of the lower pole by (4):

$$V_H = \left(\frac{32\pi i}{9w}\right)^{\frac{1}{2}} \left((H+c)^{\frac{3}{2}} - c^{\frac{3}{2}}\right).$$

Above the cloud there must be the same current i and this must be related to the potential difference between the upper pole of the cloud and the ionosphere. If u is the potential of the ionosphere with respect to earth, and if we consider the case of a cloud of positive polarity then this potential difference is $v - V - u$. We can then put:

$$v - V - u = f(i) \tag{5}$$

and we need not now specify the form of the function.

Eliminating V from (4) and (5):

$$v - u = f(i) + \left(\frac{32\pi i}{9w}\right)^{\frac{1}{2}} \left((H+c)^{\frac{3}{2}} - c^{\frac{3}{2}}\right). \tag{6}$$

Now the only effect that conditions at the earth's surface can have in (6) lies in c, which, as already stated, is a small constant. If we consider the first approximation, in which variations of c are neglected, then (6) gives one definite value of i, so that the current density is independent of the heights of the points. To the next approximation, since c is increased by lowering the points, i must be correspondingly decreased so that (6) still holds, but it is only if c becomes comparable with H that i is appreciably altered.

If the dissipation of the lower charge is almost entirely below the cloud, then the dissipating current from the lower pole must exactly balance the charge arriving from the separating process, so that the current density below the cloud is constant, quite independent of the nature of the points below.

Although the above discussion is directly applicable only to a widespread cloud, with horizontal concentrations of charge, the same conclusions are applicable without alteration to actual clouds where the charges are concentrated.

The conclusion is that for a cloud with a definite rate of separation of charge, the current density of point discharge is only very little affected by the nature of the points.

This result is so unexpected that it merits further consideration. Let us suppose that the heights of all points are reduced. Then A is decreased and M increased and so if F remained constant, i would decrease. But the reduction in the height of the points means that more lines of force can reach the ground and so F is increased. Also, if i is decreased, this decreases the space charge and so if the field near the cloud

VOL. 2 1952

is the same, fewer lines of force end on space charges, more reach the ground, increasing F, and more reach the points, increasing i again. Thus, in fact, the main effect of the reduction in height of the points is not to alter i much, but rather to increase F.

Exactly similar conclusions can be reached by discussing changes in the separation of the points.

4. Lightning Conductors

The same arguments as above can be used to discuss the effects of lightning conductors and to answer problems as old as lightning conductors themselves. SCHONLAND [8] has pointed out that two problems have arisen; first, does a lightning conductor reduce the charge in a cloud by harmless point discharge, as can be demonstrated in a laboratory experiment with a WIMSHURST machine? And, second, does the conductor increase or decrease the chance of a discharge occuring?

The first question is simply answered; we have seen that the total point discharge current below a cloud is inappreciably affected by the conditions of the points and an additional high point has practically no effect upon the dissipation of charge below a cloud. The laboratory experiment referred to does not illustrate a lightning conductor, in spite of statements that it does in certain text-books of Physics!

A lightning discharge is initiated when the field reaches, at some point, the critical value for breakdown; except for a conductor several hundred metres high, the breakdown starts in or near the cloud, where the lines of force converge on the centre of charge in the cloud. Eqs. (2), (3) and (4) refer to conditions of horizontal equipolentials and therefore cannot be exact for the conditions of localised charges. Nevertheless, it is probably legitimate to draw conclusions valid for widespread charges and then to extend them to localised charges. If the cloud charge is widespread, (2) holds and this shows that the highest field occurs just below the charge; let us consider a cloud such that this field X_H is just the limiting field for breakdowns; then we must discuss whether a lightning conductor increases or decreases X_H.

The effect of setting up a lightning conductor can be considered as producing an increase in h, the effective height of discharging points. As already remarked, this amounts to a decrease in the constant c, so that X_H is affected only to the small extent involved in c, and thus there is very little difference in the probability of discharge. So, to the first approximation, the answer to the second question is that the conductor neither increases or decreases the probability of a discharge.

If we wish, for completeness, though nor for any practical application, to go to the next approximation, the matter becomes rather more complicated. The value of X_H is given from (3):

$$X_H = \left(\frac{8\pi i}{\omega}\right)^{\frac{1}{2}} (H+c)^{\frac{1}{3}}.$$

If we consider dissipation to be almost entirely within the cloud and if $f(i)$ is small in (6), then the last term of (6) must be constant since w and u are not affected by ground conditions, so that

$$i^{\frac{1}{2}} \propto \frac{1}{[(H+c)^{\frac{3}{2}} - c^{\frac{3}{2}}]}$$

and substituting in (3)

$$X_H \propto \frac{(H+c)^{\frac{1}{2}}}{[(H+c)^{\frac{3}{2}} - c^{\frac{3}{2}}]}. \tag{7}$$

If c/H is small, this gives:

$$X_H \propto 1 - c/H$$

and

$$i \propto 1 - 3c/H.$$

On the other hand, if $f(i)$ is large compared with the last term in (6) or if dissipation takes place mainly below the cloud, i is constant as discussed earlier, and

$$X_H \propto 1 + \tfrac{1}{2} c/H.$$

The raising of the lightning conductor increases h and so decreases c. We see that this tends slightly to increase i, but its effect upon X_H depends on where the dissipation of charge occurs and no definite conclusion is possible.

The second approximation shows that a lightning conductor does serve very slightly to discharge a cloud, but we cannot determine whether it assists or hinders the initiation of a discharge.

5. Conclusion

Since the total point discharge current below a cloud is practically independent of the nature of the points below, Whipple and Scrase [3] were correct in giving equal weight to all thunderstorms in their effect upon the earth's negative charge, and in making no correction for the surface below. This conclusion can probably be extended to conditions over the sea, for the crests of waves will probably act as points for which the value of c is not comparable with H, the height of the cloud base charge. Wichmann's [5] criticism is thus not justified.

Acknowledgement—The author wishes to thank the Council of the Durham Colleges for their grant of a sabbatical year which has allowed the present work to be carried out without interruption from other duties.

VOL. 2 1952

References

[1] Wormell, T. W.; Proc. Roy. Soc. A 1927 **115** 443. [2] Schonland, B. F. J.; Proc. Roy. Soc. A 1928 **118** 252. [3] Whipple, F. J. W. and Scrage, F. J.; Geophys. Mem. 68, 1936. [4] Whipple, F. J. W.; Q.J. Roy. Met. Soc. 1938 **64** 199. [5] Wichmann, H.; Arch. für Met. Geophys. und Bioklim. 1951 **3** A 290. [6] Chalmers, J. A.; in course of publication. [7] Wilson, C. T. R.; Proc. Phys. Soc. 1925 **37** 32D. [8] Schonland, B. F. J.; The Flight of Thunderbolts. Oxford 1950.

Journal of Atmospheric and Terrestrial Physics, 1951, Vol. 2, pp. 306 to 308. Pergamon Press Ltd., London

Indices of geomagnetic activity

of the Observatories ABINGER (*Ab*), ESKDALEMUIR (*Es*) and LERWICK (*Le*)

April to June 1952

The figures given on pages 306 to 308 represent the *K*-indices for three-hour intervals, beginning with 00–03 hrs for the first and ending with 21–24 hrs for the eighth figure.

April 1952

Day	*Ab* Range for $K = 9$: 500 γ			*Es* Range for $K = 9$: 750 γ			*Le* Range for $K = 9$: 1000 γ			*Day*
	K-Indices		*Sum*	*K-Indices*		*Sum*	*K-Indices*		*Sum*	
1	4344	5442	30	4433	5432	28	6634	4431	31	1
2	6445	4455	37	5445	4455	36	6635	4677	44	2
3	5455	4655	39	5445	4655	38	6645	4466	41	3
4	3344	4565	34	4344	4454	32	6444	4466	38	4
5	4344	4645	34	4443	4545	33	5544	4556	38	5
6	6335	3455	34	5344	3345	31	5434	3346	32	6
7	4344	3454	31	4343	3344	28	5343	3344	29	7
8	5333	4444	30	4333	4344	28	5433	4344	30	8
9	4334	3445	30	4333	3435	28	5433	3435	30	9
10	5444	3233	28	5333	2223	23	7532	2123	25	10
11	3123	3221	17	2232	2221	16	2222	2221	15	11
12	3212	1133	16	3211	1132	14	3211	1122	13	12
13	3333	3232	22	3422	2131	18	4532	2232	23	13
14	2123	3423	20	1012	2423	15	2012	2323	15	14
15	3223	4332	22	3223	3331	20	2212	3332	18	15
16	3443	4433	28	3332	4332	23	3332	3332	22	16
17	2332	2112	16	2332	2211	16	3532	1221	19	17
18	1112	3445	21	0012	3345	18	1112	2345	19	18
19	3332	4445	28	3322	4444	26	5322	3434	26	19
20	2222	1323	17	2212	1213	14	2222	2213	16	20
21	3134	7865	37	3133	7865	36	3233	7777	39	21
22	4443	3432	27	4443	3332	26	4344	3322	25	22
23	3333	2332	22	3323	2332	21	3322	2331	19	23
24	1122	3424	19	1001	3424	15	1111	2434	17	24
25	1113	3312	15	2113	3311	15	2113	3311	15	25
26	1012	3342	16	1011	2332	13	1021	1332	13	26
27	1021	2123	12	1001	2112	8	2101	2112	10	27
28	4433	4433	28	3433	4333	26	3543	4444	31	28
29	3454	5666	39	3453	5666	38	5543	4678	42	29
30	5444	4555	36	6443	4555	36	7544	3557	40	30

May 1952

Day	*Ab* Range for $K = 9$: 500 γ			*Es* Range for $K = 9$: 750 γ			*Le* Range for $K = 9$: 1000 γ			Day
	K-Indices		*Sum*	*K-Indices*		*Sum*	*K-Indices*		*Sum*	
1	5445	4555	37	6444	4444	34	7545	4446	39	1
2	5434	4455	34	5534	4455	35	6634	4446	37	2
3	4433	4466	34	4433	4466	34	6642	4468	40	3
4	6333	4555	34	6433	3445	32	8533	3455	36	4
5	4334	3434	28	4333	3333	25	7333	3335	30	5
6	3343	3433	26	4343	3423	26	4444	3324	28	6
7	4344	5556	36	3344	5556	35	3444	7667	41	7
8	4343	4443	29	5433	4442	29	6533	3332	28	8
9	1121	2112	11	1011	1212	9	2111	1112	10	9
10	3123	3110	14	3112	3110	12	2212	3111	13	10
11	0212	4541	19	0122	4442	19	1112	3432	17	11
12	4332	2232	21	4322	2222	19	4322	2132	19	12
13	1122	4234	19	1121	4233	17	1112	4234	18	13
14	4221	1221	15	3221	1110	11	4321	1111	14	14
15	1121	3321	14	1001	2321	10	1112	3321	14	15
16	1111	1311	10	0111	1210	7	1111	1211	9	16
17	0111	2333	14	0001	2333	12	1111	2333	15	17
18	4344	4643	32	4333	4533	28	6432	4543	31	18
19	4545	4433	32	4444	3432	28	4634	3332	28	19
20	3333	4424	26	3333	4424	26	3333	3323	23	20
21	3333	4322	23	2223	3322	19	2223	3222	18	21
22	1122	1110	9	1112	2100	8	1121	2110	9	22
23	1113	3332	17	0112	3332	15	1111	3332	15	23
24	3223	3433	23	3222	2323	19	3222	3324	21	24
25	4333	3421	23	4233	2310	18	5332	2311	20	25
26	0045	3556	28	0135	4446	27	0134	3448	27	26
27	6344	4554	35	7535	4554	38	8645	4555	42	27
28	4334	4554	32	4323	4554	30	5323	4555	32	28
29	4444	4543	32	4534	4543	32	5544	3433	31	29
30	3332	3343	24	3432	3333	24	4432	3234	25	30
31	4432	3244	26	3431	2234	22	4442	2233	24	31

June 1952

Day	*K-Indices*		*Sum*	*K-Indices*		*Sum*	*K-Indices*		*Sum*	Day
1	2223	3332	20	2232	2332	19	3332	2222	19	1
2	2222	3332	19	1211	2331	14	2221	2231	15	2
3	2323	3333	22	2222	2233	18	2322	2233	19	3
4	4122	1222	16	3112	2221	14	3212	2222	16	4
5	2323	3322	20	2212	2322	16	2312	2322	17	5
6	3232	2111	15	2122	2111	12	3122	2111	13	6
7	1112	1231	12	1112	2121	11	1111	1131	10	7
8	4333	3544	29	3333	3433	25	3333	3333	24	8
9	3443	5544	32	3333	4433	26	3333	5433	27	9
10	4332	5434	28	3223	4434	25	3222	4334	23	10
11	3332	3433	24	2232	3433	22	4232	3423	23	11
12	3222	2332	19	3221	2222	16	3221	1221	14	12
13	0151	2232	16	0111	1123	10	0121	2122	11	13
14	1334	4555	30	1243	4554	28	1233	3445	25	14
15	3443	4333	27	3433	3333	25	2433	3323	23	15

VOL. 2 1952

June 1952 (Continued)

Day	*Ab* Range for $K = 9$: 500 γ		Sum	*Es* Range for $K = 9$: 750 γ		Sum	*Le* Range for $K = 9$: 1000 γ		Sum	Day
	K-Indices			*K-Indices*			*K-Indices*			
16	3433	4531	26	2343	3432	24	2343	3331	22	16
17	2323	3431	21	2323	3431	21	2323	3321	19	17
18	4323	3323	23	4323	3323	23	4322	3223	21	18
19	2222	3312	17	2111	3221	13	1112	3221	13	19
20	2222	2221	15	2111	2221	12	1111	2221	11	20
21	1112	2322	14	1112	2212	12	1112	2222	13	21
22	1233	5544	27	1133	5444	25	1233	5444	26	22
23	3344	5434	30	3344	4433	28	5533	4324	29	23
24	4533	4433	29	4433	4433	28	6623	5543	34	24
25	1123	3443	21	2112	3333	18	3112	4333	20	25
26	4334	4422	26	4323	4421	23	4323	3421	22	26
27	2332	3433	23	2222	3333	20	2332	3333	22	27
28	3123	3311	17	2113	3221	15	3213	2211	15	28
29	1122	2245	19	1112	2234	16	2112	2236	19	29
30	5765	3322	33	5766	3222	33	5985	3211	34	30

Book review

Latin Treatises on Comets between 1238 and 1368 A.D.; Edited by Lynn Thorndike.. 274 pp, University of Chicago Press, 1951, $ 5.00; Cambridge University Press, 37/6.

This learned compilation makes us aware of the comparatively small increase in our knowledge of the origin of comets during the last thousand years! Actually the century and a half covered by Dr. THORNDIKE marked a regression to the views of ARISTOTLE who held that comets were terrestrial exhalations, bursting into flame when they reached the upper air. The commentaries of ALBERTUS MAGNUS and THOMAS AQUINAS on ARISTOTLE's "Meteorology" are, in fact, largely in defence of ARISTOTLE's views. No doubt ALBERTUS MAGNUS was fully justified in demolishing the views of JOHN OF DAMASCUS for whom comets were formed as signs signifying the death of kings "by divine command to suit the time and again are dissolved." But history does not support his attack on SENECA who stated his opinion as follows:

> "A comet is a star created together with the works of nature, whose nature it may be is ignored, yet that it is not transient fire is proved thus. Whatever things air produces are short-lived. How then can a comet last long in air, when the air itself is not long permanent? For they are born in what is flighty and mobile, and it cannot be that fixed fire should have its seat and so pertinaciously adhere in a wandering body . . . A comet has its seat and so is not quickly expelled but traces its measured course, nor is extinguished but passes on . . ."

It is unfortunate that ALBERTUS found so many objections, not only to SENECA's views, but also to the opinions of other "moderns." How much these commentaries of ALBERTUS and AQUINAS influenced the subsequent writers of the 13th and 14th centuries is impossible to say, but in the event, the belief in ARISTOTLE's opinion is very solid.

In addition to these two commentaries (which are in translation), the book contains the Latin text of nine other treatises on comets and one on the solar eclipse of 1337. Six of the treatises deal with the specific comets of 1299, 1301, 1315, 1337, 1338, and 1368; the other three are more general. In each case Dr. THORNDIKE provides an explanatory introduction and commentary. The book makes pleasant reading and is a refreshing change from contemporary scientific papers on comets.

A. C. B. LOVELL

Journal of Atmospheric and Terrestrial Physics, 1952 ,Vol. 2, pp. 309 to 323. Pergamon Press Ltd., London

Nitrogen in the upper Atmosphere

S. Deb

Institute of Radio Physics and Electronics, Calcutta University

(Received 16 April 1952)

ABSTRACT

An attempt has been made to find out from theoretical considerations the probable distribution of atomic nitrogen in the upper atmosphere. Two mechanisms of production of atomic nitrogen are considered: (i) predissociation of N_2 molecules by absorption in the Lyman-Birge-Hopfield band within the range 1150–1250 Å as suggested by Herzberg & Herzberg, and, (ii) dissociative recombination of N_2^+ ions and electrons as suggested by Mitra. It is shown that the N atoms thus produced will disappear largely through two-body radiative recombination process. The method of calculation employed for determining the distribution is a modification of the one adopted by Penndorf in estimating the dissociation of molecular oxygen in the upper atmosphere. The results show that while Herzberg & Herzbergs mechanism is operative within the range of heights 70–170 km, that of Mitra is effective in the much higher range—200–400 km. The dissociation in the former region is found to be less than 1% and the distribution of N atoms is found to have a sharp maximum round the height of 90 km. The dissociation in the latter region, on the other hand, is found to increase from 5% at 170 km to about 25% at 400 km, the distribution not showing any maximum (the rise is at first sharp; this is followed by a very slow but continuous increase with height). Calculations also show that the total number of N atoms in a vertical column of one cm² cross-section is about 10^{16}.

Estimate has also been made of the abundance of N_2^+ ions in the upper atmosphere. It is shown that although these are largely removed through the dissociative recombination process, a sufficient number is still left to account for the observed twilight flash of negative bands of N_2^+ ions.

The probable time required for the attainment of equilibrium of the various processes involved in the calculation is also considered. It is found that the concentration of N_2^+ ions reaches an equilibrium value very quickly, but it requires years for the concentration of N atoms to reach the same.

VOL.
2
1952

1. Introduction

The identification of the forbidden lines λ 3466 [^2P—^4S] and λ 5198 [^2D—^4S] in the auroral spectrum proves conclusively the presence of nitrogen in the atomic state in the upper atmosphere, in the auroral regions at least. It is of considerable interest to enquire if nitrogen, in such state, is also present in regions other than auroral. This may be possible if there be dissociation of N_2 through the absorption of extreme ultraviolet solar radiation. Unfortunately, no dissociation process by direct absorption of solar ultraviolet radiation, as in the case of O_2, is known for N_2. However, two processes have recently been suggested—one by Herzberg & Herzberg [11] and the other by Mitra [14] by which ultraviolet light absorption may ultimately lead to dissociation of N_2 molecules.

According to Herzberg & Herzberg [11] there may be predissociation of nitrogen molecules by absorption in the Lyman-Birge-Hopfield band within the range 1150–1250 Å.

According to Mitra [14] dissociative recombination of N_2^+ ion and electron may produce excited nitrogen atoms.

It is the purpose of this paper to investigate the probable distribution of atomic nitrogen in the upper atmosphere by assuming that both of the above processes are operative during daylight hours.

2. Basic Reactions and Working Formulae

(i) *Production of atomic nitrogen*

For the process suggested by Herzberg & Herzberg, N_2 molecules as are excited to $v' = 6$–9 of the $a\,^1\Pi_g$ state by absorption in the range 1150–1250 Å, may dissociate by radiationless transition. Thus, we have

$$N_2 + h\nu \xrightarrow{f} N_2(a\,^1\Pi_g) \xrightarrow{g} N + N. \tag{1}$$

Here f is the amount of solar energy absorbed per N_2-molecule per unit time (at the atmospheric height under consideration) within the active waveband and g is the fraction of the excited $N_2\,(a\,^1\Pi_g)$ molecules produced per second which undergo dissociation. The sun is assumed to emit like a black body at 6000° K.

Note. If the dissociation energy of nitrogen molecules $D(N_2)$ be assigned the older value of 7·38 eV then one of the N atoms is in the normal state (^4S) and the other in the excited state (^2D). If, however, Gaydon's [6] value of 9·76 eV is accepted then both the atoms are in the normal state (^4S).

In the process suggested by Mitra the primary reaction is the production of N_2^+ ions by absorption in the range $\lambda < 795$ Å. Thus

$$N_2 + h\nu \xrightarrow{f'} N_2^+(X') + e. \tag{2}$$

Here f' is the amount of solar energy in the spectral range $\lambda < 795$ Å absorbed per N_2 molecule per unit time (at the height under consideration). The nitrogen ions thus formed may undergo dissociative recombination as follows:

$$N_2^+(X') + e \xrightarrow{\beta} N(^2P) + N(^2D), \tag{3}$$

where β is the coefficient of dissociative recombination of N_2^+ ions and electrons.

We note that in this reaction the energy released by the left hand side is 15·58 eV (first ionization potential of N_2) and that required on the right hand side is the dissociation energy of N_2 (9·76 eV) plus energies of excitation of the two N atoms, 3·56 eV and 2·37 eV respectively, *i.e.* total of 15·69 eV. The demand is, therefore, greater than the supply by 0·11 eV. However, it has been pointed out by Mitra that if the N_2^+ (X') ion be in any of the low vibrational levels ($v'' = 1$, 2 etc.) then the slight defect is made up. Alternatively, according to Bates (private communication to Mitra) the deficiency may be made up if the recombining electron is sufficiently energetic.

Note. Since the intensity of the solar radiation decreases by absorption it is obvious that the value of f will also decrease with decrease of height. For a wavelength range $\lambda_1 \to \lambda_2$, its value, at a particular height, is given by

$$f = J_{h\lambda} \cos\chi\, B_\lambda\, d\lambda, \tag{4}$$

where, $J_{h\lambda}\, d\lambda$ is the number of photons entering the earth's atmosphere per cm² at the given height within the wavelength range $\lambda \to \lambda + d\lambda$, χ is the solar zenith angle and B_λ is the absorption cross-section for the wavelength λ [12].

It is easily seen that

$$J_{h\lambda}\, d\lambda = J_{\infty\lambda} \exp\left(-\int_{\infty}^{h} B_\lambda [N_2]_h\, dh\right) d\lambda, \tag{5}$$

where, $J_{\infty\lambda}\, d\lambda$ is the number of photons within the range $\lambda \to \lambda + d\lambda$ entering per cm² from outside the atmosphere and $[N_2]_h$ is the number density of N_2-molecules at the height h.

Hence, if B_λ is assumed to be constant over the absorption band under consideration, then one obtains from [4] and [5]

$$f = B \exp\Big(- B \int_{\infty}^{h} [N_2]_h \, dh \Big) \cos\chi \int_{\lambda_1}^{\lambda_2} J_{\infty\lambda} \, d\lambda \tag{6}$$

$$= B \exp(- B \Phi) \, 3 \times 10^{17} \cos\chi \int_{x_1}^{x_2} \frac{x^2}{\exp(x) - 1} , \tag{7}$$

where,

$$\Phi = \int_{\infty}^{h} [N_2]_h \, dh \tag{8}$$

and

$$x = \frac{h c}{\lambda k T} . \tag{9}$$

(ii) *Disappearance of atomic nitrogen*

The equilibrium link between the number densities of molecular and atomic nitrogen is provided by recombination of the N atoms. This may occur through a three-body collision process,

$$N + N + M \xrightarrow{k'} M' + N_2, \tag{10}$$

or, through a two-body process

$$N + N \xrightarrow{\alpha} N_2 + h\nu, \tag{11}$$

where k' and α are the coefficients of the three-body and two-body recombination processes respectively.

The k'-value for N atoms is not known. But, from a consideration of the k'-values for three-body recombination of other atoms (*e.g.* oxygen atoms) it may be supposed that for N atoms k' is less than 10^{-36} cm^6/s. Hence, for the purpose of our calculations this will be taken as the upper limit of the value of k'.

The probable value of α for the two-body recombination of the N atoms is discussed in Sec. 3 (vii) to follow. It is shown that if the energy of dissociation of N_2 molecules be 9·76 eV as advocated by GAYDON, then the value of α will lie within the range 10^{-18}–10^{-21} cm^3/s.

Now, to determine which of the two processes of disappearance of N atoms—two-body or three-body—will be effective, we first note that

$$\frac{\text{probability of recombination by three-body collision}}{\text{probability of recombination by two-body collision}} = \frac{(1 - \Theta) \, n \, k'}{\alpha} ,$$

where n is the concentration of the total of the different kinds of particles and Θ is the ratio of the concentration of recombining particles to n. Hence, with the above assumed values of k' and α, it easily follows that the three-body process cannot dominate above a height where the number density of the particles is 10^{15}/cm^3 *i.e.* above 80 km. It will, however, be shown that both of the regions of interest—that of predissociation by the HERZBERG & HERZBERGs process and that of dissociative recombination by MITRAs process—are almost wholly above 80 km. Hence, in considering the disappearance of N atoms we need take into account only the two-body recombination process.

VOL. 2 1952

(iii) *Working formulae*

(a) HERZBERG & HERZBERGs *predissociation process*—The rate of production of N atoms by predissociation as given by Eq. (1) is

$$2fg\,[N_2],$$

where $[N_2]$ is the concentration of N_2 molecules.

The rate of disappearance of N atoms by two-body recombination as given by Eq. (11) is

$$2\alpha\,[N]^2,$$

where $[N]$ is the concentration of N atoms.

Thus,

$$\frac{d\,[N]}{dt} = 2fg\,[N_2] - 2\alpha\,[N]^2. \tag{12}$$

The equilibrium concentration $[N]_{eq}$ of N atoms is, therefore, given by

$$[N]_{eq} = \sqrt{\frac{fg\,[N_2]}{\alpha}}, \tag{13}$$

provided only a small fraction of the N_2 molecules is dissociated.

(b) MITRAs *dissociative recombination process*—For this case we have,

rate of production of N_2^+ ions $= f'\,[N_2]$,
rate of disappearance of N_2^+ ions $= \beta\,[N_2^+]\,[e]$,
rate of production of N atoms $= 2\beta\,[N_2^+]\,[e]$,
rate of disappearance of N atoms $= 2\alpha\,[N]^2$,

where $[N_2^+]$ and $[e]$ are the concentrations of N_2^+ ions and electrons respectively. Hence,

$$\frac{d\,[N_2^+]}{dt} = f'\,[N_2] - \beta\,[N_2^+]\,[e], \tag{14}$$

and the concentration of N_2^+ under equilibrium condition is given by

$$[N_2^+]_{eq} = \frac{f'\,[N_2]}{\beta\,[e]}, \tag{15}$$

provided only a small fraction of N_2 molecules is ionised.

When Eq. (15) is valid we have,

$$\frac{d\,[N]}{dt} = 2f'\,[N_2] - 2\alpha\,[N]^2, \tag{16}$$

and the concentration of N atoms under equilibrium condition is given by

$$[N]_{eq} = \sqrt{\frac{f'\,[N_2]}{\alpha}}. \tag{17}$$

3. NUMERICAL VALUES ASSUMED FOR THE REACTION CONSTANTS

Inspection of Eqs. (12) to (17) shows that to make quantitative estimates of the rates of the various processes of production and disappearance of N_2^+ ions and N atoms, as also the equilibrium values of the same, we require a knowledge of the following: (i) the active wave bands, (ii) the values of absorption cross sections of N_2 for the active processes, (iii) height distribution of neutral N_2 molecules, (iv) value of g, (v) value of the dissociative recombination coefficient of N_2,

(vi) height distribution of electrons and (vii) two-body radiative recombination coefficient of N atoms.

We give below the most probable values of these as assumed in the calculation and the factors influencing their choice.

(i) *Active wave bands*—For HERZBERG & HERZBERGs process the active wave band is 1150–1250 Å.

For the process suggested by MITRA there is absorption in two wavebands, one $\lambda < 795$ Å (leading to the production of N_2^+ ions in the normal X' state) and the other at $\lambda < 661$ Å (leading to the production of N_2^+ ions in the excited A' state). However, since the energy density is much higher for the longer waveband we need consider only this band. Further, as oxygen is known to have an absorption continuum at 744 Å, we will restrict our consideration over the wavelength range 795–750 Å.

(ii) *Values of absorption cross-section of N_2*—Absorption by N_2 molecule in the range 1150–1250 Å is very weak—about 40 cm of air at S.T.P. being required for complete absorption [13]. On this basis, the absorption cross-section is likely to be much less than 10^{-20} cm^2 throughout the band. We will, however, work with the value 10^{-20} cm^2 remembering that this will give an upper limit of the value of percentage dissociation and of the height of the region of maximum dissociation. For simplifying the calculation we shall further assume that the value of the cross-section is constant within the active waveband.

The absorption cross-section at the limiting value of 795 Å for the ionization process will be assumed to be 10^{-17} cm^2 after BATES and MASSEY [2]. For this case also we shall assume the value to be constant within the active waveband.

(iii) *Height distribution of neutral N_2-molecules*—It is to be noted that the purpose of the calculations that follow is to determine the height distribution of N atoms or, what comes to the same thing, that of N_2 molecules as affected by the dissociative processes. However, for calculations under equilibrium condition (with which we are concerned) it is obvious that to obtain the intensity of the dissociating radiation at any level we require a knowledge—at least an approximate one—of the total number of absorbing N_2 molecules above this level. This may appear as begging the question at the start. But, if it is found that the finally calculated height distribution does not differ widely from the distribution assumed at the start, then we shall be justified in concluding that the calculated results conform to the actual condition of distribution. We, therefore, start with the distribution of N_2 molecules above 100 km as suggested from various considerations by GERSON [8], *viz.*, very little dissociation below 100 km and dissociation increasing from nil at 100 km to 25% at 300 km. The values of N_2 concentrations at different levels thus obtained are shown in column 2 of Tables 1 and 2.

(iv) *Value of g*—g is the ratio of the probability of disappearance of an excited $^1\Pi_g$ molecule through predissociation to that through all the other processes. Since the probability is inversely related to the life time, we may also put

$$g = \frac{\tau_0}{\tau'}$$

where τ_0 is the mean life time of the excited state as given by the reciprocal of the sum of the probabilities of all possible transitions and τ' is the mean life until

VOL. 2 1952

radiationless dissociation. We can attempt at making an estimate of the value of g from estimates of those of τ_0 and τ'.

Consider first the various processes of disappearance of the excited $^1\Pi_g(v' > 6)$ molecules. These are: (i) collisional de-excitation, (ii) spontaneous transition to the ground state accompanied by radiation, (iii) stimulated transition to the ground state accompanied by radiation, (iv) stimulated transition to the higher states $^1\Sigma_u$ and $^1\Pi_u$ by absorption of solar energy, and finally (v) radiationless transition, namely, pre-dissociation. In the equilibrium condition the total of the rates of all these processes of disappearance must equal the rate of production of the excited $^1\Pi_g(v' < 6)$ molecules. Let us now consider the relative importances of these various processes. Process (i) may be the most rapid one at high and at moderate pressures, but its effect will be small at low pressures. Processes (ii), (iii) and (v) involve forbidden transitions and, as such, are necessarily very slow [10], [11]. Process (iv), however, involves allowed transitions and hence is important. GAYDON and HERMAN [5–7] have found that the intensities of some of these transitions in emission *i.e.* transition from higher to lower states, are comparable to that of the second positive system which is one of the prominent systems in the visible and near ultraviolet spectrum of N_2. Thus, the probability of the reverse processes, namely, transition from lower to the higher state by absorption, must also be high. Further, the stimulating radiation required lies in the range 2300–3200 Å, for which the available solar energy density is high. One may conclude, therefore, that at all pressures where collision plays a negligible part, the excited $^1\Pi_g$ molecules will disappear largely through process (iv). It therefore follows that the value of τ_0 will be determined by the transition probability of this—the quickest process. Let us try to make an estimate of this probability. We know that the free life time of an excited state, for which there are allowed transitions to some lower states, is of the order 10^{-8} to 10^{-9} s. Hence, we may assume that the EINSTEIN A-probability for the spontaneous transition $^1\Sigma_u \rightarrow {}^1\Pi_g$ is of the order 10^8. The EINSTEIN B-probability for transition by stimulation in the average of the wavelength range 2000–3000 Å is, therefore, of the order 10^{18} (since $A/B = 10^{10}$). Now, the average density of the solar radiation in this wavelength range (taking account of the dilution factor) is about 10^{-10} erg. Hence the probability of the transition $^1\Pi_g \rightarrow {}^1\Sigma_u$ by absorption of solar energy in this wavelength range, is $10^{-10} \times 10^{18} = 10^8$. τ_0 may therefore be assumed to be of the order 10^{-8} s.

Regarding the value of τ' one may make a guess from the following consideration. Bands with predissociation characteristics, namely, sudden break off of the series have been observed in this case only in emission [3]. This suggests that the probability of predissociation of an excited molecule is at best of the same order as that of its spontaneous transition to the ground state [6], [9]. In other words, the minimum value that τ' can have is of the same order as the free liefe time τ of the excited state. The value of τ, although not accurately known, may be roughly inferred from the following expression due to SAHA and KOTHARI [15] for the life of a metastable atom,

$$\tau = \frac{3\, c^2\, m^2}{\pi\, e^2\, h\, \bar{\nu}^3}, \tag{20}$$

where $\bar{\nu}$ is the wave number of the forbidden line, c the velocity of light, e and m the charge and mass of electron. For $\lambda = 1200$ Å, the above relation gives

$\tau = 10^{-4}$ s, but, this may be too low an estimate. The upper limit of the value of τ', the mean life until radiationless dissociation, may therefore also be taken as 10^{-4} s. Hence, we have for the upper limit of the value of $g = 10^{-8}/10^{-4} = 10^{-4}$.

We, therefore, assume g to be 10^{-4} in regions of low pressure. For intermediate heights where collisional frequency is greater than 10^4/s the value of g is likely to be smaller. This condition exists in the range of heights up to the F_1 region and we assume that $g = 10^{-5}$ within the range 70–170 km.

(v) *Value of dissociative recombination coefficient* β—As already pointed out, for the dissociative recombination to proceed according to Eq. (3), either the reacting N_2^+ (X') ions must be in a vibrationally excited state or the reacting electron must be energetic enough to supply the deficiency of 0·11 eV. Of the two, it is possible that the first condition is not fully satisfied when, as assumed, the N_2 molecule is ionized at its first ionization potential. But, the second condition is undoubtedly satisfied, because, according to all estimates, the F-region has a high temperature $> 1000°$ K [8] and, for such temperature, the average kinetic energy of the electrons corresponds to 0·1 eV. Hence, assuming the dissociative recombination process to be operative, its rate coefficient may be assumed to be 10^{-7} cm^3/s after BATES [1]. The coefficient is, however, pressure dependent, and, for the pressures as exist above 200 km the value may be somewhat lower—say by about one order. For simplicity we shall assume a constant value of β, 10^{-8} cm^3/s throughout the range 100–400 km.

(vi) *Height distribution of electrons*—This may be calculated by assuming a parabolic distribution in the regions of maximum ionization. The necessary data are readily available from ionospheric records. A set of values obtained in this manner is shown in column 4 of Table 3.

(vii) *Radiative recombination coefficient* α—The value of α for N atoms is not known; but, a rough guess of the order of the value may be made from the following consideration. As already mentioned the probability of photodissociation of N_2 molecule is much smaller than that of O_2 molecules. It, therefore, follows from the principle of detailed balancing that the probability of the reverse process, namely, radiative recombination of two N-atoms is also much less than that for two oxygen atoms. Now, the value of α for two normal oxygen atoms is 10^{-20} cm^3/s. We shall, therefore, assume that the corresponding value for two normal N atoms is 10^{-21} cm^3/s. It has, however, been pointed out by TA-YOU WU [17] that if one of the combining N atoms be in an excited state (*e.g.* ^2D state) the probability is much increased. For such case a reasonable assumption is 10^{-18}–10^{-19} cm^3/s.

Now, for HERZBERG & HERZBERGs predissociation process the atoms produced are in the normal state (provided one accepts GAYDONs value of dissociation energy, viz. 9·76 eV). Hence, when considering this process we assume the value of α to be 10^{-21} cm^3/s. (Some of the normal N atoms produced may be excited by light absorption to the ^2D state. But this number will be small as, the state being a metastable one, the absorption coefficient will be very small. According to DUFAY [4] the relative abundance of such excited N atoms cannot exceed 1 in 10^6.)

For the dissociative recombination process as suggested by MITRA one of the N atoms produced is in the ^2D state and the other in the ^2P state. In view of

VOL. 2 1952

the fact that the excited ^2D state has very long life there will be considerable concentration of these excited atoms throughout day and night. Hence, when considering this process we assume α to be 10^{-18}–10^{-19} cm^3/s.

4. Numerical Calculations and Results

The calculations are carried out in the following steps:

(i) The distribution of N_2 molecule concentration with height is plotted on a graph paper with the values given in column 2 of Tables 1 and 2.

Table 1

Height (km)	$[N_2]$	Φ	f	g	$[N]_{eq}$
1	2	3	4	5	6
170	8×10^{10}	$5{\cdot}13 \times 10^{17}$	$4{\cdot}8 \times 10^{-9}$	10^{-5}	$1{\cdot}96 \times 10^{9}$
160	$9{\cdot}8 \times 10^{10}$	$5{\cdot}98 \times 10^{17}$	$4{\cdot}8 \times 10^{-9}$	10^{-5}	$2{\cdot}16 \times 10^{9}$
150	$2{\cdot}6 \times 10^{11}$	$7{\cdot}77 \times 10^{17}$	$4{\cdot}78 \times 10^{-9}$	10^{-5}	$3{\cdot}5 \times 10^{9}$
140	$4{\cdot}9 \times 10^{11}$	$1{\cdot}15 \times 10^{18}$	$4{\cdot}75 \times 10^{-9}$	10^{-5}	$4{\cdot}82 \times 10^{9}$
130	8×10^{11}	$1{\cdot}79 \times 10^{18}$	$4{\cdot}7 \times 10^{-9}$	10^{-5}	$6{\cdot}14 \times 10^{9}$
120	$2{\cdot}6 \times 10^{12}$	$3{\cdot}49 \times 10^{18}$	$4{\cdot}66 \times 10^{-9}$	10^{-5}	$1{\cdot}1 \times 10^{10}$
110	7×10^{12}	$8{\cdot}29 \times 10^{18}$	$4{\cdot}42 \times 10^{-9}$	10^{-5}	$1{\cdot}76 \times 10^{10}$
100	$2{\cdot}5 \times 10^{13}$	$2{\cdot}43 \times 10^{19}$	$3{\cdot}79 \times 10^{-9}$	10^{-5}	$3{\cdot}08 \times 10^{10}$
90	$8{\cdot}8 \times 10^{13}$	$8{\cdot}08 \times 10^{19}$	$2{\cdot}12 \times 10^{-9}$	10^{-5}	$4{\cdot}34 \times 10^{10}$
80	$6{\cdot}8 \times 10^{14}$	$4{\cdot}68 \times 10^{20}$	$4{\cdot}32 \times 10^{-11}$	10^{-5}	$1{\cdot}7 \times 10^{10}$
70	5×10^{15}	$1{\cdot}47 \times 10^{21}$	$4{\cdot}13 \times 10^{-16}$	10^{-5}	$1{\cdot}4 \times 10^{8}$

Table 2

Height (km)	$[N_2]$	Φ	f'	$[N]_{eq}$	$[N]_{eq}$
1	2	3	4	5	6
400	$8{\cdot}8 \times 10^{9}$	0	$2{\cdot}089 \times 10^{-10}$	$1{\cdot}36 \times 10^{9}$	$4{\cdot}3 \times 10^{9}$
390	$8{\cdot}88 \times 10^{9}$	$8{\cdot}84 \times 10^{15}$	$1{\cdot}9 \times 10^{-10}$	$1{\cdot}3 \times 10^{9}$	$4{\cdot}12 \times 10^{9}$
380	$8{\cdot}95 \times 10^{9}$	$1{\cdot}77 \times 10^{16}$	$1{\cdot}75 \times 10^{-10}$	$1{\cdot}25 \times 10^{9}$	$3{\cdot}95 \times 10^{9}$
370	9×10^{9}	$2{\cdot}67 \times 10^{16}$	$1{\cdot}59 \times 10^{-10}$	$1{\cdot}2 \times 10^{9}$	$3{\cdot}8 \times 10^{9}$
360	$9{\cdot}07 \times 10^{9}$	$3{\cdot}57 \times 10^{16}$	$1{\cdot}46 \times 10^{-10}$	$1{\cdot}15 \times 10^{9}$	$3{\cdot}63 \times 10^{9}$
350	$9{\cdot}15 \times 10^{9}$	$4{\cdot}48 \times 10^{16}$	$1{\cdot}33 \times 10^{-10}$	$1{\cdot}11 \times 10^{9}$	$3{\cdot}51 \times 10^{9}$
340	$9{\cdot}3 \times 10^{9}$	$5{\cdot}4 \times 10^{16}$	$1{\cdot}22 \times 10^{-10}$	$1{\cdot}06 \times 10^{9}$	$3{\cdot}35 \times 10^{9}$
330	$9{\cdot}6 \times 10^{9}$	$6{\cdot}34 \times 10^{16}$	$1{\cdot}1 \times 10^{-10}$	$1{\cdot}03 \times 10^{9}$	$3{\cdot}25 \times 10^{9}$
320	$9{\cdot}9 \times 10^{9}$	$7{\cdot}31 \times 10^{16}$	10^{-10}	$9{\cdot}95 \times 10^{8}$	$3{\cdot}14 \times 10^{9}$
310	10^{10}	$8{\cdot}3 \times 10^{16}$	$9{\cdot}1 \times 10^{-11}$	$9{\cdot}54 \times 10^{8}$	$3{\cdot}02 \times 10^{9}$
300	$1{\cdot}1 \times 10^{10}$	$9{\cdot}35 \times 10^{16}$	$8{\cdot}24 \times 10^{-11}$	$9{\cdot}52 \times 10^{8}$	$3{\cdot}01 \times 10^{9}$
290	$1{\cdot}25 \times 10^{10}$	$1{\cdot}05 \times 10^{17}$	$7{\cdot}28 \times 10^{-11}$	$9{\cdot}54 \times 10^{8}$	$3{\cdot}02 \times 10^{9}$
280	$1{\cdot}45 \times 10^{10}$	$1{\cdot}18 \times 10^{17}$	$6{\cdot}42 \times 10^{-11}$	$9{\cdot}65 \times 10^{8}$	$3{\cdot}05 \times 10^{9}$
270	$1{\cdot}6 \times 10^{10}$	$1{\cdot}33 \times 10^{17}$	$5{\cdot}52 \times 10^{-11}$	$9{\cdot}42 \times 10^{8}$	$2{\cdot}98 \times 10^{9}$
260	$1{\cdot}85 \times 10^{10}$	$1{\cdot}5 \times 10^{17}$	$4{\cdot}61 \times 10^{-11}$	$9{\cdot}23 \times 10^{8}$	$2{\cdot}91 \times 10^{9}$
250	$2{\cdot}15 \times 10^{10}$	$1{\cdot}7 \times 10^{17}$	$3{\cdot}82 \times 10^{-11}$	$8{\cdot}97 \times 10^{8}$	$2{\cdot}82 \times 10^{9}$
240	$2{\cdot}6 \times 10^{10}$	$1{\cdot}94 \times 10^{17}$	3×10^{-11}	$8{\cdot}83 \times 10^{8}$	$2{\cdot}79 \times 10^{9}$
230	$2{\cdot}9 \times 10^{10}$	$2{\cdot}21 \times 10^{17}$	$2{\cdot}29 \times 10^{-11}$	$8{\cdot}15 \times 10^{8}$	$2{\cdot}58 \times 10^{9}$
220	$3{\cdot}3 \times 10^{10}$	$2{\cdot}5 \times 10^{17}$	$1{\cdot}68 \times 10^{-11}$	$7{\cdot}45 \times 10^{8}$	$2{\cdot}36 \times 10^{9}$
210	4×10^{10}	$2{\cdot}88 \times 10^{17}$	$1{\cdot}17 \times 10^{-11}$	$6{\cdot}84 \times 10^{8}$	$2{\cdot}17 \times 10^{9}$
200	$4{\cdot}6 \times 10^{10}$	$3{\cdot}31 \times 10^{17}$	$7{\cdot}63 \times 10^{-12}$	$5{\cdot}92 \times 10^{8}$	$1{\cdot}88 \times 10^{9}$
190	$5{\cdot}3 \times 10^{10}$	$3{\cdot}8 \times 10^{17}$	$4{\cdot}68 \times 10^{-12}$	$4{\cdot}98 \times 10^{8}$	$1{\cdot}77 \times 10^{9}$
180	$6{\cdot}3 \times 10^{10}$	$4{\cdot}38 \times 10^{17}$	$2{\cdot}62 \times 10^{-12}$	$4{\cdot}06 \times 10^{8}$	$1{\cdot}28 \times 10^{9}$
170	8×10^{10}	$5{\cdot}09 \times 10^{17}$	$1{\cdot}41 \times 10^{-12}$	$3{\cdot}36 \times 10^{8}$	$1{\cdot}06 \times 10^{9}$

(ii) The values of Φ, the total number of N_2 molecules above any height, are now calculated with the help of this graph using trapezoidal rule of approximate integration with 10 km intervals (contributions from regions above 400 km are neglected). The calculated values are shown in column 3 of Tables 1 and 2.

(iii) The values of f, the amount of solar energy absorbed per N_2 molecule per second, for different heights are next calculated with the help of Eq. (7) using the tabulated values of Φ and the data as assumed in Secs. 3 (i) and 3 (ii) for $\chi = 0$ *i.e.* for condition at noon at equinoxes. In Table 1, column 4 are shown the values as obtained for the HERZBERG & HERZBERGs process. In Table 2, column 4, are shown the values as obtained for MITRAs process.

(iv) The equilibrium concentration values of N at different heights are calculated with the help of Eqs. (13) and (17) utilising the tabulated values of f and the data as assumed in Secs. 3 (iv) and 3 (vii). The calculated values are shown in column 5 of Table 1 for $\alpha = 10^{-21}$ cm^3/s, in column 5 of Table 2 for $\alpha = 10^{-18}$ cm^3/s and in column 6 of the same Table for $\alpha = 10^{-19}$ cm^3/s. The values are also depicted graphically in Fig. 1 and in curve 1 of Fig. 2. (The lower scale of the abscissa of Fig. 2 is for the case $\alpha = 10^{-18}$ cm^3/s and the upper one for the case $\alpha = 10^{-19}$ cm^3/s.)

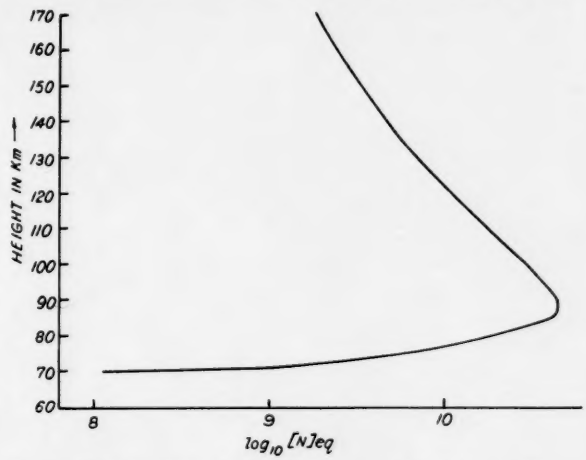


Fig. 1. Illustrating the distribution with height, under equilibrium condition, of atomic nitrogen, $[N]_{eq}$, produced by predissociation of N_2 molecules according to HERZBERG & HERZBERG'S process.

Table 3

Height (km)	$[N_2]$	f'	$[e]$	$[N_2^+]_{eq}$
1	2	3	4	5
170	8×10^{10}	$1{\cdot}41 \times 10^{-12}$	$6{\cdot}24 \times 10^{4}$	$1{\cdot}85 \times 10^{2}$
200	$4{\cdot}6 \times 10^{10}$	$7{\cdot}63 \times 10^{-12}$	$2{\cdot}2 \times 10^{5}$	$1{\cdot}59 \times 10^{2}$
220	$3{\cdot}3 \times 10^{10}$	$1{\cdot}68 \times 10^{-11}$	$2{\cdot}5 \times 10^{5}$	$2{\cdot}22 \times 10^{2}$
240	$2{\cdot}6 \times 10^{10}$	3×10^{-11}	$2{\cdot}2 \times 10^{5}$	$3{\cdot}44 \times 10^{2}$
270	$1{\cdot}6 \times 10^{10}$	$5{\cdot}52 \times 10^{-11}$	$9{\cdot}5 \times 10^{5}$	$9{\cdot}3 \times 10^{1}$
300	$1{\cdot}1 \times 10^{10}$	$8{\cdot}24 \times 10^{-11}$	10^{6}	$9{\cdot}1 \times 10^{1}$
340	$9{\cdot}3 \times 10^{9}$	$1{\cdot}22 \times 10^{-10}$	$9{\cdot}1 \times 10^{5}$	$1{\cdot}23 \times 10^{2}$
400	$8{\cdot}8 \times 10^{9}$	$2{\cdot}089 \times 10^{-10}$	5×10^{5}	$3{\cdot}7 \times 10^{2}$

(v) The equilibrium values of N_2^+ ions are calculated for a few typical heights using the values of β as given in Sec. 3 (v), the values of electron concentration as tabulated in column 4 of Table 3 and the values of N_2 molecule concentrations as tabulated in column 2 of Table 3. The results are shown in column 5 of Table 3.

VOL. 2 1952

(vi) Next, all the above calculations are repeated taking into account the attenuation of solar energy due to absorption by the N_2 molecules existing in the region above 400 km. Unfortunately, the height distribution of N_2 molecules above 400 km is not known. A rough estimate is, however, made by assuming

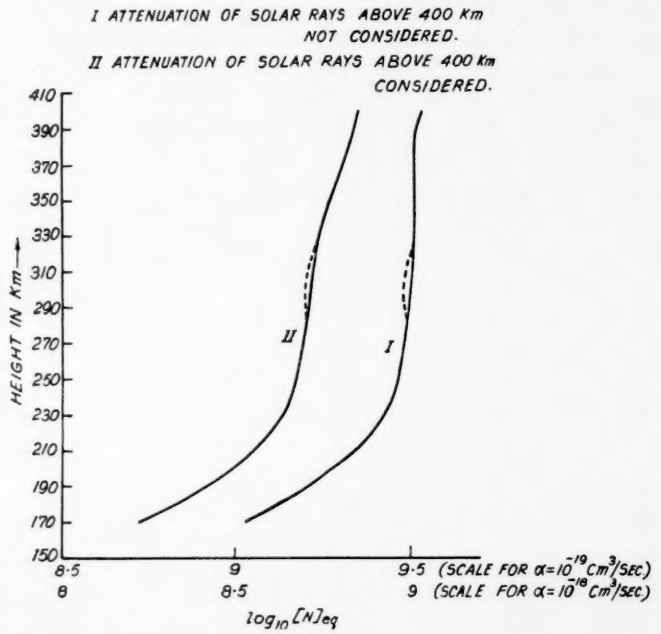


Fig. 2. Illustrating the distribution with height, under equilibrium condition, of atomic nitrogen $[N]_{eq}$ produced by dissociative recombination of N_2^+ ions and electrons as suggested by MITRA. In curve II the total attenuation of the active solar radiation, in reaching the level under consideration, has been taken into account. In curve I the attenuation above 400 km has been neglected. *Note:* The calculated values of $[N]_{eq}$ shows anomalous cusp round 300 km as shown by dashes. This is probably due to different nature of the assumptions made for calculation above and below the 300 km level. However, the cusp has been ignored and the curve has been drawn following the general trend.

that the scale height (H) remains constant above this level. Under such condition the value of Φ at 400 km is given by

$$\int_0^\infty N_0 \exp\left(-\frac{h}{H}\right) dh,$$

where, N_0 is the concentration of N_2 molecules at the 400 km level. Consideration of the intensity distribution along the length of the auroral streamers indicate that the value of H is about 146 km at 400 km [8]. With this value of H and from the distribution of N_2 molecules as assumed in Sec. 3 (iii), the values of Φ at different heights are calculated. These new values of Φ are shown in column 2 of Tables 4 and 5. The corresponding values of f for $\chi = 0$ are given in column 3 of these tables. The new values of equilibrium concentration of N atoms are shown in column 5 of Table 4 for $\alpha = 10^{-21}$ cm^3/s, in column 4 of Table 5 for $\alpha = 10^{-18}$ cm^3/s,

and in column 5 of the same table for $\alpha = 10^{-19}$ cm^3/s. These results are also depicted graphically in Fig. 2, curve II.

Table 4

Height (km)	Φ	f'	g	$[N]_{eq}$
1	2	3	4	5
170	$6{\cdot}42 \times 10^{17}$	$4{\cdot}8 \times 10^{-9}$	10^{-5}	$1{\cdot}96 \times 10^{9}$
160	$7{\cdot}27 \times 10^{17}$	$4{\cdot}79 \times 10^{-9}$	10^{-5}	$2{\cdot}15 \times 10^{9}$
150	$9{\cdot}06 \times 10^{17}$	$4{\cdot}77 \times 10^{-9}$	10^{-5}	$3{\cdot}49 \times 10^{9}$
140	$1{\cdot}28 \times 10^{18}$	$4{\cdot}74 \times 10^{-9}$	10^{-5}	$4{\cdot}8 \times 10^{9}$
130	$1{\cdot}92 \times 10^{18}$	$4{\cdot}71 \times 10^{-9}$	10^{-5}	$6{\cdot}12 \times 10^{9}$
120	$3{\cdot}62 \times 10^{18}$	$4{\cdot}61 \times 10^{-9}$	10^{-5}	$1{\cdot}07 \times 10^{10}$
110	$8{\cdot}42 \times 10^{18}$	$4{\cdot}42 \times 10^{-9}$	10^{-5}	$1{\cdot}76 \times 10^{10}$
100	$2{\cdot}44 \times 10^{19}$	$3{\cdot}79 \times 10^{-9}$	10^{-5}	$3{\cdot}08 \times 10^{10}$
90	$8{\cdot}09 \times 10^{19}$	$2{\cdot}1 \times 10^{-9}$	10^{-5}	$4{\cdot}3 \times 10^{10}$
80	$4{\cdot}68 \times 10^{20}$	$4{\cdot}32 \times 10^{-11}$	10^{-5}	$1{\cdot}7 \times 10^{10}$
70	$1{\cdot}47 \times 10^{21}$	$4{\cdot}13 \times 10^{-16}$	10^{-5}	$1{\cdot}4 \times 10^{8}$

Table 5

Height (km)	Φ	f'	$[N]_{eq}$	$[N]_{eq}$
1	2	3	4	5
400	$1{\cdot}29 \times 10^{17}$	$5{\cdot}72 \times 10^{-11}$	$7{\cdot}01 \times 10^{8}$	$2{\cdot}22 \times 10^{9}$
390	$1{\cdot}37 \times 10^{17}$	$5{\cdot}26 \times 10^{-11}$	$6{\cdot}83 \times 10^{8}$	$2{\cdot}17 \times 10^{9}$
380	$1{\cdot}46 \times 10^{17}$	$4{\cdot}9 \times 10^{-11}$	$6{\cdot}62 \times 10^{8}$	$2{\cdot}10 \times 10^{9}$
370	$1{\cdot}55 \times 10^{17}$	$4{\cdot}4 \times 10^{-11}$	$6{\cdot}3 \times 10^{8}$	$1{\cdot}99 \times 10^{9}$
360	$1{\cdot}64 \times 10^{17}$	4×10^{-11}	6×10^{8}	$1{\cdot}9 \times 10^{9}$
350	$1{\cdot}73 \times 10^{17}$	$3{\cdot}68 \times 10^{-11}$	$5{\cdot}91 \times 10^{8}$	$1{\cdot}87 \times 10^{9}$
340	$1{\cdot}82 \times 10^{17}$	$3{\cdot}55 \times 10^{-11}$	$5{\cdot}5 \times 10^{8}$	$1{\cdot}77 \times 10^{9}$
330	$1{\cdot}91 \times 10^{17}$	$3{\cdot}06 \times 10^{-11}$	$5{\cdot}43 \times 10^{8}$	$1{\cdot}72 \times 10^{9}$
320	$2{\cdot}01 \times 10^{17}$	$2{\cdot}77 \times 10^{-11}$	$5{\cdot}25 \times 10^{8}$	$1{\cdot}66 \times 10^{9}$
310	$2{\cdot}11 \times 10^{17}$	$2{\cdot}51 \times 10^{-11}$	5×10^{8}	$1{\cdot}58 \times 10^{9}$
300	$2{\cdot}22 \times 10^{17}$	$2{\cdot}25 \times 10^{-11}$	$4{\cdot}99 \times 10^{8}$	$1{\cdot}58 \times 10^{9}$
290	$2{\cdot}34 \times 10^{17}$	$1{\cdot}99 \times 10^{-11}$	5×10^{8}	$1{\cdot}58 \times 10^{9}$
280	$2{\cdot}48 \times 10^{17}$	$1{\cdot}72 \times 10^{-11}$	$5{\cdot}1 \times 10^{8}$	$1{\cdot}61 \times 10^{9}$
270	$2{\cdot}63 \times 10^{17}$	$1{\cdot}48 \times 10^{-11}$	$4{\cdot}88 \times 10^{8}$	$1{\cdot}54 \times 10^{9}$
260	$2{\cdot}80 \times 10^{17}$	$1{\cdot}24 \times 10^{-11}$	$4{\cdot}78 \times 10^{8}$	$1{\cdot}51 \times 10^{9}$
250	$3{\cdot}00 \times 10^{17}$	$1{\cdot}0 \times 10^{-11}$	$4{\cdot}65 \times 10^{8}$	$1{\cdot}47 \times 10^{9}$
240	$3{\cdot}24 \times 10^{17}$	$8{\cdot}2 \times 10^{-12}$	$4{\cdot}6 \times 10^{8}$	$1{\cdot}45 \times 10^{9}$
230	$3{\cdot}52 \times 10^{17}$	$6{\cdot}2 \times 10^{-12}$	$4{\cdot}24 \times 10^{8}$	$1{\cdot}34 \times 10^{9}$
220	$3{\cdot}83 \times 10^{17}$	$4{\cdot}54 \times 10^{-12}$	$3{\cdot}88 \times 10^{8}$	$1{\cdot}22 \times 10^{9}$
210	$4{\cdot}2 \times 10^{17}$	$3{\cdot}13 \times 10^{-12}$	$3{\cdot}54 \times 10^{8}$	$1{\cdot}11 \times 10^{9}$
200	$4{\cdot}63 \times 10^{17}$	$2{\cdot}03 \times 10^{-12}$	$3{\cdot}06 \times 10^{8}$	$9{\cdot}5 \times 10^{8}$
190	$5{\cdot}13 \times 10^{17}$	$1{\cdot}24 \times 10^{-12}$	$2{\cdot}57 \times 10^{8}$	$8{\cdot}13 \times 10^{8}$
180	$5{\cdot}71 \times 10^{17}$	$6{\cdot}8 \times 10^{-13}$	$2{\cdot}07 \times 10^{8}$	$6{\cdot}5 \times 10^{8}$
170	$6{\cdot}42 \times 10^{17}$	$3{\cdot}36 \times 10^{-13}$	$1{\cdot}64 \times 10^{8}$	$5{\cdot}18 \times 10^{8}$

Note. In Tables 1 and 4 the values of f and $[N]_{eq}$ are shown for HERZBERG & HERZBERGs process. In Tables 2 and 5 these quantities are shown for MITRAs process. Calculations based on HERZBERG & HERZBERGs process are not carried out above 170 km, because, above this height the contribution due to MITRAs process becomes predominent. For regions below 170 km MITRAs process plays a minor rôle and hence this process is left out of consideration.

VOL. 2 1952

5. DISCUSSION

Inspection of Figs. 1 and 2 shows that the two suggested processes for the dissociation of N_2 molecules—predissociation (HERZBERG & HERZBERG) and dissociative recombination of N_2^+ ions (MITRA), are operative at quite different levels. The former is effective in the 80–100 km region near the *E* layer of the ionosphere and the latter above 170 km in the *F* region.

For the predissociation process, the fraction of the N_2 molecules dissociated is extremely small. The distribution of N atoms shows a pronounced maximum in the 80–90 km region (Fig. 1). The results are practically the same whether

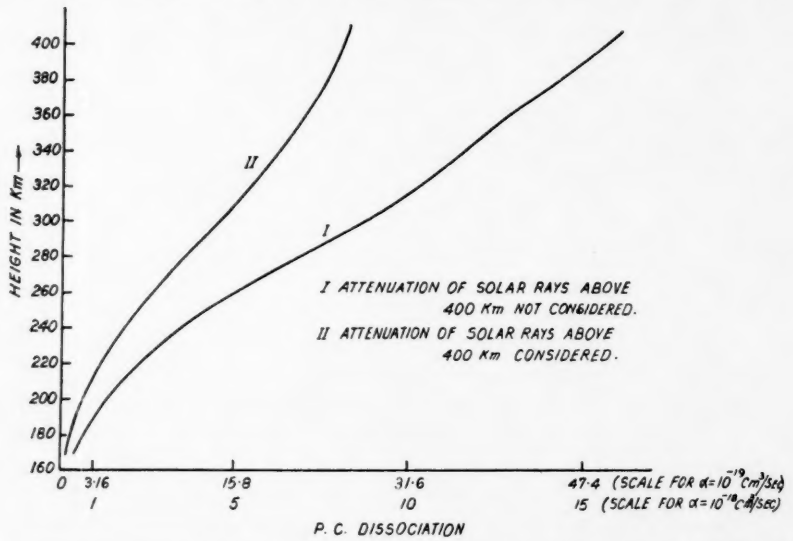


Fig. 3. Calculated percentage dissociation of N_2 molecules with height. In curve II the total attenuation of the active solar radiation, in reaching the level under consideration, has been taken into account. In curve I the attentuation above 400 km has been neglected.

or not one takes into account absorption by the nitrogen molecules above 400 km. This, of course, is as expected. As already pointed out, the absorption cross-section of N_2 molecules assumed for the calculation is on the higher side. The actual maximum of the distribution, therefore, lies in a region lower down by a few kilometres. The value of the percentage dissociation at this level will, however, be smaller. But this may be higher if the value of g is assumed to be larger or that of α smaller.

For the dissociative recombination process, the fraction of N_2-molecules dissociated becomes appreciable only above 200 km (Fig. 3). For $\alpha = 10^{-19}$ cm^3/s, the dissociation is about 50% at 400 km, if the attenuation of solar rays above this height is ignored. The corresponding value for $\alpha = 10^{-18}$ cm^3/s is 15·5%. If the attenuation above 400 km is taken into account, the percentage dissociation decreases to 25 for $\alpha = 10^{-19}$ cm^3/s and to 8 for $\alpha = 10^{-18}$ cm^3/s. It is to be recalled that in computing the values of Φ it was assumed that the dissociation increased from 0 to 25 per cent within the range 100–400 km. The results obtained

for the region above 300 km, therefore, agree fairly well with this assumption. For lower regions the agreement appears to be unsatisfactory. For example, on the basis of the initial assumption, the dissociation at 170 km should be about 6 per cent; whereas, MITRAs process gives the value of only 0·5. However, at this height, contribution to the percentage of dissociation by HERZBERG & HERZBERGs process is about 2·5, if $g = 10^{-5}$ and 7·8 if $g = 10^{-4}$. A linear increase in the dissociation from 0 to 25 per cent over the range 100–400 km appears, therefore, to be a reasonable conclusion. A better agreement between the assumed and the calculated values is of little significance on account of the many uncertainties in the magnitudes of the parameters involved.

It is to be noted that according to the results obtained there is no maximum in the height distribution of the N atoms in the higher regions. This is because of the very slow decrease of N_2 molecular concentration with height on account of the rising temperature assumed in this region. This compensates the effect of increase of f with height, which would otherwise produce a maximum. The slow and continuous increase of $[N]_{eq}$ with height is, therefore, as expected. It follows from Tables 2 and 5 that the total number of N atoms in a column of cross-section 1 sq cm is of the order of 10^{16}. This agrees well with a recent estimate by DUFAY [4] made from observations of the intensity of $^4S-^2D$ line of N atoms in the twilight flash spectrum in summer.

VOL. 2 1952

It is interesting to note that an additional consequence of MITRAs process is the disappearance of N_2^+ ions. As will be seen from Table 3, column 5, the number density of N_2^+ ions under equilibrium condition, if MITRAs process is operative, is low. However, in the higher regions where β, the coefficient of dissociative recombination becomes small [1] sufficient number of such ions will still be left to account for the observed twilight flash of the negative bands of N_2^+ ions when the high atmosphere is illuminated by solar rays.

Before concluding it will be interesting to discuss the time required for the attainment of equilibrium for the two processes of atomic nitrogen production.

Consider first the production and disappearance of N_2^+ ions. One readily obtains from Eq. (14).

$$[N_2^+] = [N_2^+]\,(1 - \exp(-\beta[e]\,t)), \tag{21}$$

provided the concentration of N_2 molecules remains approximately unaltered during the progress of the reaction. It follows from Eq. (21) that the time required for the attainment of equilibrium will be of the order of

$$\frac{1}{\beta[e]} \simeq 100 - 1000 \text{ seconds}.$$

Next, let us consider the increase in the concentration of N atoms. It follows from Eqs. (12) and (16) that the time t required for the concentration to build up to a value $[N]_t$ is

$$t = \frac{1}{2\sqrt{XY}} \ln \frac{\sqrt{\frac{X}{Y}} + [N]_t}{\sqrt{\frac{X}{Y}} - [N]_t},$$

where, $Y = 2\alpha$

$X = 2gf[N_2]$ for HERZBERG & HERZBERGs process

$= 2f'[N_2]$ for MITRAs process.

The time required for $[N]_t$ to grow to a value $\frac{e-1}{e}[N]_{eq}$ is given approximately by

$$\frac{1}{\sqrt{X\,Y}} = \frac{1}{2\sqrt{f'\alpha\,[N_2]}}$$

for MITRAs process and,

$$= \frac{1}{2\sqrt{g f \alpha\,[N_2]}}$$

for HERZBERG & HERZBERGs process.

Thus, the time for the attainment of equilibrium for MITRAs process would be of the order of 10^9 s and that for HERZBERG & HERZBERGs process 10^{10} s, *i.e.* a matter of years for both the processes. It thus appears that unless the recombination rate at night is extremely small, equilibrium may not be attained at all. When the active radiation is withdrawn, we have to put the first term on the righthand side of Eqs. (12) and (16) equal to zero. This gives the time t for the concentration to fall to $[N]_t$ from its equilibrium value

$$t = \frac{[N]_{eq} - [N]_t}{2\alpha\,[N]_t\,[N]_{eq}}.$$

This relation shows that for the concentration of N atoms to become small

$t \simeq 10^{18}$ seconds, for MITRAs process.

$t \simeq 10^{20}$ seconds, for HERZBERG & HERZBERGs process.

Recombination during night will, therefore, be extremely small.

6. CONCLUDING REMARKS

From the results of calculations in the preceding sections it is seen that molecular nitrogen would be dissociated in significant proportion in the upper atmosphere both by predissociation as suggested by HERZBERG & HERZBERG and by dissociative recombination as suggested by MITRA. The calculations are, however, based on many uncertain data—the weakest link in the chain of calculations being the value of g, the probability of radiationless dissociation of the excited $N_2(a\,^1\Pi_g)$ molecules in relation to their probability of removal from the excited state by other processes. Available evidences, however, do not suggest that the actual value of g would be greatly different from the one assumed in the above calculations. It is also to be noted that even if the value of g is assumed to be increased by an order (*i.e.* $g = 10^{-4}$–10^{-3} instead of 10^{-5}–10^{-4}) the resulting equilibrium value of N atom concentration would be increased from 0·5 to about 1·6 per cent at 120 km and from 2·2 to about 7 per cent at 160 km. Nevertheless, the numerical results obtained are to be regarded more as indicating the rough order of dissociation of N_2 than as giving actual quantitative values.

It may be noted that according to the calculations given at the end of Sec. 5 the equilibrium state is attained very slowly. Hence, for attaining thermodynamic equilibrium, the individual atoms and molecules will have to remain in the region under consideration for a considerably long time. This condition is, however, hardly realised in the so-called exosphere, *viz.*, the region extending above 300 km. According to SPITZER [16] the average length of time that atoms and molecules spend in this region is 3–4 hrs which is extremely small compared to the time

for attainment of equilibrium. It is, therefore, very likely that thermodynamic equilibrium is never attained in this region and the actual concentration of N atoms at any height there is considerably smaller than that given in Tables 2 and 5.

Acknowledgement—The author wishes to thank Professor S. K. MITRA for suggesting the problem and for giving valuable guidance and help in course of the preparation of the paper.

REFERENCES

[1] BATES, D. R.; Phys. Rev. 1950 **78** 492. [2] BATES, D. R. and MASSEY, H. S. W.; Proc. Roy. Soc. A 1946 **187** 261. [3] DOUGLAS, A. E. & HERZBERG, G.; Can. J. Phys. 1951 **29** 294. [4] DUFAY, M.; Comptes Rendus 1951 **232** 419. [5] GAYDON, A. G.; Proc. Roy. Soc. A 1944 **182** 286. [6] GAYDON, A. G.; Dissociation Energies and Spectra of Diatomic Molecules, Chapman and Hall, London 1947. [7] GAYDON, A. G. and HERMAN, R.; Proc. Phys. Soc. 1946 **58** 292. [8] GERSON, N. C.; Rep. on. Progr. in Phys. 1951 **14** 316. [9] HERZBERG, G.; Molecular Spectra and Molecular Structure, p. 437. Prentice Hall, New York 1939. [10] HERZBERG, G.; Phys. Rev. 1946 **69** 362. [11] HERZBERG, G. & HERZBERG, L.; Nature, Lond. 1948 **161** 283. [12] PENNDORF, R.; J. Geophys. Res. 1949 **54** 7. [13] MITRA, S. K.; The Upper Atmosphere (2nd edition). The Asiatic Society, Calcutta 1952. [14] MITRA, S. K.; Nature, Lond. 1951 **167** 897. [15] SAHA, M. N. & KOTHARI, D. S.; Naturwiss. 1929 **17** 271. [16] SPITZER, L.; The Atmosphere of the Earth and Planets (edited by P. KUIPER). University of Chicago Press, Illinois 1949. [17] TA-YOU WU; Phys. Rev. 1944 **66** 65.

VOL. 2 1952

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 324 to 336. Pergamon Press Ltd., London

Theoretical group heights of reflection of 150 kc/s radio waves vertically incident on the ionosphere*

NORMAN DAVIDS

Ionosphere Research Laboratory, The Pennsylvania State College, State College, Pennsylvania

(Received 22 April 1952)

ABSTRACT

This paper is concerned with the theoretical determination of the group heights of reflection of long radio waves vertically incident on the ionosphere. The theoretically expected group heights for 250 μsec Gaussian shaped pulses are determined utilizing: (1) a wave theory treatment including coupling, (2) a CHAPMAN-like ionospheric model including all variables and their height variations and (3) a fundamental radio frequency or carrier component of 150 kc/s.

From the theoretical viewpoint the continuous wave (c.w.) results of GIBBONS and NERTNEY [1], [2] are extended by including the dispersive characteristics of the above model. This is accomplished by determining the received pulse characteristics by means of a response function developed from a suitable Fourier-Hermite series which includes the frequency dependent effects of the model. It is found that the rapid changes in polarization near the lower edge of the *E*-region, called the "coupling region", are very effective in generating reflected waves or pulses under suitable conditions, *i.e.* late night hours associated with low *E*-layer critical frequencies. However, the dispersive characteristics are small so that almost negligible group retardation is to be expected. At a higher level in the layer rapid variations in the index of refraction for one of the characteristic wave components occur in what is called the "reflection region". Here the time-delays are large for low critical frequency models.

Some geometrical discussion is included concerning polarization and coupling factors. If properly considered, the ordinary and extraordinary modes can be given a representation (in a certain complex space) as orthogonal principal directions.

The above theory is compared with the experimentally observed group heights on a diurnal basis. Good agreement is obtained and is further strengthened by comparison with some recent measurements of R. E. JONES on the diurnal change in phase (c.w.) paths.

1. INTRODUCTION

Enough group-height data at 150 kc/s have now been obtained, both at this Laboratory and elsewhere, to establish the existence of a triplet of first-hop echoes. These so-called triple splits can be observed especially at night when the *E*-layer critical frequency is small. The theory here has been treated for the c.w. case in [1].

The first-hop echo of least delay, for the frequency of interest, is presumed due to the "coupling layer" where rapid changes in the polarizations of the magneto-ionic components occur as a function of height.

As described in [1] and [2], the first-hop echo of intermediate delay is identified with the *E*-layer reflection of the ordinary wave, made up both of the reflected ordinary component and the extraordinary forward scattered wave generated in the coupling region. This echo is usually observed throughout both day and night hours, *i.e.* at high and low *E*-layer critical frequencies.

Under certain conditions a first-hop echo of still greater delay, of the order of 150 km, is frequently noted. While this is currently being studied from both

* The research reported in this paper has been supported, in part, by the Geophysical Research Division of the Air Forces Cambridge Research Center under Contract AF 19(122)-4.

theoretical and experimental viewpoints, no satisfactory theory is as yet available and is not considered further in this paper.

In [1] and [2] some of the first-order coupling effects for c.w. at 150 kc/s were calculated for an ionosphere model of CHAPMAN form in which the E-layer critical frequency ranged from 0·55 to 4·4 mc/s. This paper takes as its basis a similar model. The theoretical analysis of pulse transmission is made by using the Fourier Integral theorem. Such analyses were made by RYDBECK [6] for square waves. For pulses of the order of 10 μsec it is only necessary to study a 10 kc/s band centering on the carrier frequency. Major attention needs only to be confined to the 0·55 mc/s critical frequency night-time E-layer model since, it has been shown, coupling effects are most dominant near this value. Some additional effects not present in the c.w. solutions are discussed, especially time delays in the E-layer. The normal pulse echoes observed here show little distortion of shape, as is shown theoretically, but the time delays are large enough to affect the interpretation of group-height data. This effect is shown to be accentuated if shorter pulse lengths than those currently used by this Laboratory are considered.

2. The Isotropic Directions

We shall use the notation of RYDBECK [3] except that variation with time in the steady state is taken in the form $e^{j\omega t}$ instead of $e^{-j\omega t}$. If rectangular coordinates are chosen so that Z is measured vertically upward, Y in the direction of magnetic north, X in that of magnetic east, then the components of the electric polarization vector $\boldsymbol{P}$ are connected with those of the electric field $\boldsymbol{E}$ by the relations (see [3])

$$\left.\begin{aligned} \frac{-E_x}{4\pi} &= \chi_0^2\left(1-j\frac{\nu}{\omega}\right)P_x - j\gamma_L P_y + j\gamma_T P_z \\ \frac{-E_y}{4\pi} &= j\gamma_L P_x + \chi_0^2\left(1-j\frac{\nu}{\omega}\right)P_y + \chi_0^2\left(1-j\frac{\nu}{\omega}\right)P_z \\ \frac{-E_z}{4\pi} &= -j\gamma_T P_x \end{aligned}\right\} \tag{2.1}$$

where N is the number of electrons per cm^3, $\chi_0 = \omega/\omega_c$, $\gamma_L = \chi_0^2\,(\omega_H/\omega)\cos\theta_P$, $\gamma_T = \chi_0^2\,(\omega_H/\omega)\sin\theta_P$, $\omega_c = \sqrt{4\pi Ne^2/m}$ is the critical frequency in sec^{-1}, θ_P the angle between the Earth's field and the Z-axis.

Assuming further that no net free charge is present and that all quantities vary with Z only, the divergence equation reduces to div $\boldsymbol{D} = \partial D_z/\partial z = 0$. If we neglect the presence of constant electric fields, we have $D_z = E_z + 4\pi P_z = 0$ by means of which the number of independent components in (2.1) reduces to two.

The condition

$$\frac{E_x}{P_x} = \frac{E_y}{P_y} \tag{2.2}$$

defines two principal states of propagation having the polarizations

$$\frac{E_y^{(1)}}{E_x^{(1)}} = u_1; \qquad \frac{E_y^{(2)}}{E_x^{(2)}} = u_2 \tag{2.3}$$

which satisfy the relation

$$u_1\,u_2 = 1 \tag{2.4}$$

This leads to an interesting geometric relationship between the transverse projections of $\boldsymbol{E}^{(1)}$ and $\boldsymbol{E}^{(2)}$. Denote these respectively by $\tilde{\boldsymbol{E}}^{(1)}$ and $\tilde{\boldsymbol{E}}^{(2)}$. From (2.2) it follows

that $\tilde{\boldsymbol{E}}^{(i)}$ and $\tilde{\boldsymbol{P}}^{(i)}$ $(i = 1, 2)$ have the same directions—thus the aforementioned principal modes define directions of isotropy in the x, y-plane. The polarization numbers u_1, u_2 as defined by (2.3), represent the tangents of these direction angles—in general, complex.

As suggested in a recent Letter to the Editor [4] let us imagine an imaginary space so that the direction of $\boldsymbol{E}$ is given by

$$\tan\theta = j\frac{E_y}{E_x} = j\,u; \tag{2.5}$$

As in real space, the direction of $\tilde{\boldsymbol{E}}$ and $\tilde{\boldsymbol{P}}$ coincide when $u = u_i$ $(i = 1, 2)$. For these directions we have, by (2.4)

$$\tan\theta_1 \tan\theta_2 = (j\,u_1)(j\,u_2) = -1\,. \tag{2.6}$$

Thus the principal modes become orthogonal in this space. This property permits the use of these directions as a set of oblique coordinate axes. If we identify suscripts 1 and 2 with the conventional ordinary and extraordinary modes and let $\boldsymbol{i}_0, \boldsymbol{i}_e$ denote unit vectors in these directions, respectively, then the equations of transformation of coordinates are given by

$$\tilde{\boldsymbol{E}} = E_0\,\bar{\boldsymbol{i}}_0 + E_e\,\bar{\boldsymbol{i}}_e$$

where

$$E_0 = E_x\cos\theta + j\,E_y\sin\theta = \frac{(E_x - u_1 E_y)}{\sqrt{1-u_1^2}}$$

$$E_e = -E_x\sin\theta + j\,E_y\cos\theta = j\frac{(-u_1 E_x + E_y)}{\sqrt{1-u_1^2}}.$$

In terms of the polarization u of the given field vector we have

$$E_0 = E_x\frac{(1 - u\,u_1)}{\sqrt{1-u_1^2}} \tag{2.7}$$

$$E_e = j\,E_x\frac{(u - u_1)}{\sqrt{1-u_1^2}}. \tag{2.8}$$

With these resolution formulas the effect of an input wave of arbitrary polarization can be determined from those of pure o- and e-waves. For these waves the above formulas reduce respectively to

$$\left.\begin{aligned} &E_0^{(1)} = E_x^{(1)}\sqrt{1-u_1^2}, && E_e^{(1)} = 0 \\ &E_0^{(2)} = 0\,, && E_e^{(2)} = j\,E_x\frac{\sqrt{1-u_1^2}}{u_1} = E_x^{(2)}\sqrt{1-u_2^2}. \end{aligned}\right\} \tag{2.9}$$

In terms of the π-waves used in [3], where

$$\Pi_1 = E_x^{(1)}\sqrt{1-u_1^2}; \quad \cdot \quad \Pi_2 = E_y^{(2)}\sqrt{1-u_1^2}, \tag{2.10}$$

we have the relations

$$E_0^{(1)} = \Pi_1; \qquad E_e^{(2)} = \Pi_2.$$

3. Dispersion in the Coupling Region

The oblique axes of the last section can be used to derive the coupled wave equations. Physically the coupling between the principal modes arises from the variation in the imaginary space of the principal axes. If we consider the vector $\boldsymbol{\Pi}_1$, its absolute

derivative is given by

$$\frac{d\boldsymbol{\Pi}_1}{dz} = \Pi_1' \boldsymbol{i}_0 + \Pi_1 \theta' \boldsymbol{i}_e$$

where θ is the complex angle as defined by (2.5), and the primes denote the (scalar) derivative with respect to z. By repeated application of the same operation, the absolute acceleration is given by

$$\begin{aligned}\frac{d^2\boldsymbol{\Pi}_1}{dz^2} &= d\frac{(\Pi_1'\boldsymbol{i}_0)}{dz} + d\frac{(\Pi_1\theta'\boldsymbol{i}_e)}{dz} \\ &= \Pi_1''\boldsymbol{i}_0 + \Pi_1'\theta'\boldsymbol{i}_e + \Pi_1'\theta'\boldsymbol{i}_e + \Pi_1\theta''\boldsymbol{i}_e - \Pi_1\theta'^2\boldsymbol{i}_0. \end{aligned} \tag{3.1}$$

To express θ' in terms of the polarizations, we differentiate the relation (2.5) with respect to z,

$$\theta' \sec^2\theta = j\,u'$$

or

$$\theta' = \frac{j\,u'}{(1-u^2)} = j\,M$$

in the notation of GIBBONS and NERTNEY [1]. This function may be considered as a measure of coupling strength. For all practical purposes the coupling region is where this function (and its derivative) are appreciable. We may say that M is the space rate of twist (in radians per km) of the principal polarizations, as a result of which each principal mode induces a wave of the opposite mode. Eq. (3.1) now becomes

$$\frac{d^2\boldsymbol{\Pi}_1}{dz^2} = (\Pi_1'' + M^2\Pi_1)\boldsymbol{i}_0 + j\,(\Pi_1 M' + 2\Pi_1' M)\boldsymbol{i}_e. \tag{3.2}$$

Similarly we have an equation for the Π_2-component:

$$\frac{d^2\boldsymbol{\Pi}_2}{dz^2} = (\Pi_2'' + M^2\Pi_2)\boldsymbol{i}_e - j\,(\Pi_2 M' + 2\Pi_2' M)\boldsymbol{i}_0. \tag{3.3}$$

If we consider the $\boldsymbol{\Pi}$-vectors are representing the oscillations of particles, then the acceleration must be balanced by that caused by the restoring forces of the system; consequently, adding (3.2) and j times (3.3),

$$\frac{d^2(\Pi_1 + j\,\Pi_2)}{dz^2} + k_0^2\,\varepsilon_1\,\Pi_1 + j\,k_0^2\,\varepsilon_2\,\Pi_2 = 0. \tag{3.4}$$

Written out in full

$$\begin{aligned}(\Pi_1'' + M^2\Pi_1 + \Pi_2 M' + 2\Pi_2' M)\boldsymbol{i}_0 + j\,(\Pi_2'' + M^2\Pi_2 + \Pi_1 M' + 2\Pi_1' M)\boldsymbol{i}_e + \\ + k_0^2\,\varepsilon_1\,\Pi_1\boldsymbol{i}_0 + j\,k_0^2\,\varepsilon_2\,\Pi_2\boldsymbol{i}_e = 0.\end{aligned}$$

Equating coefficients of $\boldsymbol{i}_0$ and $\boldsymbol{i}_e$, we obtain

$$\left.\begin{aligned}\Pi_1'' + (k_0^2\,\varepsilon_1 + M^2)\,\Pi_1 &= -\Pi_2 M' - 2\Pi_2' M \\ \Pi_2'' + (k_0^2\,\varepsilon_2 + M^2)\,\Pi_2 &= -\Pi_1 M' - 2\Pi_1' M\end{aligned}\right\} \tag{3.5}$$

These are the coupled wave-equations. They have been treated differently in [1] and [3]. The eq. (3.4) is a vector formulation.

The terms in the above equations can be interpreted as forces in a rotating coordinate system. $M^2\Pi_1$ is the centrifugal force, $\Pi_2 M'$ the tangential force, and $2\Pi_2' M$ the Coriolis force.

It should be noted that actually, for the E-layer models of interest to us, M and M' are negligible everywhere except over a rather narrow interval of height somewhat below the E-layer maximum here assumed constant at 115 km. The indices of refraction $\sqrt{\varepsilon_1}$, $\sqrt{\varepsilon_2}$ are slowly-varying parameters close to their free-space values in this region so that coupling effects arise almost solely through the variation of M. These effects may be classified into two parts: (1) induction of one characteristic mode by the other due to the terms on the right side of the coupled wave eqs. (3.1), (2) variations of the "effective dielectric constant".

$$k_0^2 \varepsilon_i + M^2 \qquad i = 1, 2$$

also appearing in the same equations. If M^2 varies rapidly this factor would lead to additional reflections. In the present paper we shall confine ourselves to the study of strictly coupling effects under type (1).

In setting up the boundary conditions for the coupling region we may use in the adjacent regions the "WKB solutions" appropriate to a slowly-varying medium. In [1] these are used to solve the uncoupled equations, whose solutions are then cycled back to give a first-order solution to the coupled equation. It is shown that the effect of the coupling is to excite a new wave (forward-scattered) in the direction of propagation of the incident wave and also a back-scattered wave in the reverse direction. In terms of a unit amplitude upgoing ordinary wave these waves will both be extraordinary, and are given by

$$\Pi_2^{\pm} = n_2^{-\frac{1}{2}} \exp\left(\mp j k_0 \int_0^z n_2 \, d\zeta\right) F^{\pm} \tag{3.6}$$

where the upper signs refer to the upgoing wave, the lower signs to the downgoing wave, $N_2 = \sqrt{\varepsilon_2}$, and

$$F^{+} = -\int_a^b M \exp\left(-j k_0 \int_0^z (n_1 - n_2) \, d\zeta\right) dz \tag{3.7}$$

$$F^{-} = \tfrac{1}{2} \int_a^b M (n_1 n_2)^{-\frac{1}{2}} (n_2 - n_1) \exp\left(-j k_0 \int_0^z (n_1 + n_2) \, d\zeta\right) dz. \tag{3.8}$$

A numerical evaluation of these integrals was made for a normalized Chapman distribution with maximum density at a constant height of 115 km, and for a critical frequency $f_c = 0{\cdot}55$ mc/s. This frequency was chosen since it corresponds approximately to late night-time conditions when coupling echoes are strongest. The 150 kc/s case originally evaluated by GIBBONS and NERTNEY was extended to 140 and 160 kc/s. For the forward scattered wave amplitudes at these frequencies we have the following values (see Table 1). The back scattered wave amplitudes are of the order of 1% and hence negligible by comparison. Fig. 1 shows the variation of the coupling factor M for these frequencies.

Table 1. Forward scattered wave amplitudes

f	F^+
140 kc/s	$0{\cdot}334\, e^{-j52°}$
150 kc/s	$0{\cdot}318\, e^{-j65°}$
160 kc/s	$0{\cdot}397\, e^{-j35°}$

A method for using this information for evaluating pulse transmission will be developed below in the next section.

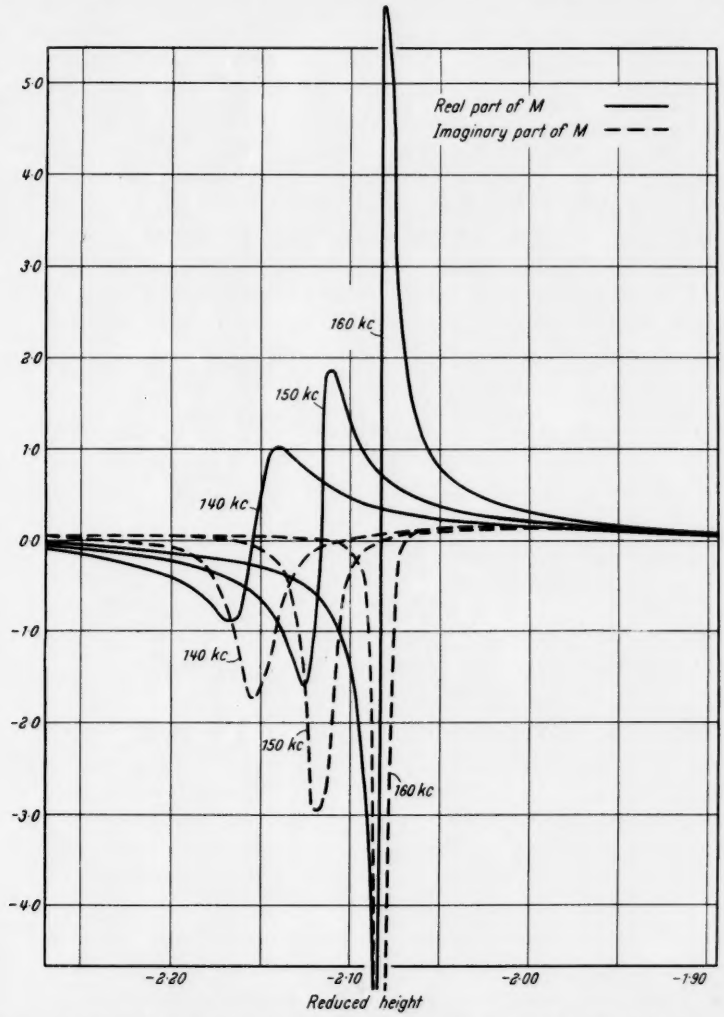


Fig. 1. Coupling factor M over coupling region for 0·55 mc/s critical frequency.

4. Pulse-Transmission Analysis

From the elementary plane electromagnetic wave solution $\exp j(\omega t - k_0 n z)$ we may construct, by superposition, the solution

$$f(z,t) = \int_a^b A(\omega) \exp j(\omega t - k_0 n z)\, d\omega \tag{4.1}$$

where $A(\omega)$ is an arbitrary function of frequency ω, subject only to the condition that the integral converge. We have here a boundary value problem where we have prescribed

$$f(0,t) = f(t) = \int_a^b A(\omega) \exp j\omega t\, d\omega \tag{4.2}$$

VOL. 2 1952

where $f(t)$ represents the form of the transmitted pulse as it leaves the ground. It will be assumed in this analysis that our pulse is "pure Gaussian" in shape and appears as the envelope of a modulated carrier at frequency $\omega = \omega_0$, centred about $t = 0$:

$$f(t) = A \exp\left[-\frac{t^2}{T^2} + j\,\omega_0 t\right] \tag{4.3}$$

where T is a measure of semi-pulse-length (from maximum amplitude A to amplitude A/e) and ω_0 is the carrier angular frequency. The pulse envelope is significant only if T includes sufficiently many wavelengths of the carrier.

A harmonic analysis has been made of our pulses by NERTNEY, and it is expected that the pure Gaussian pulse is a reasonably accurate representation. The Fourier transform or frequency spectrum of $f(t)$ is then given by

$$g(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)\, e^{-j\omega t}\, dt \tag{4.4}$$

$$= \frac{AT}{\sqrt{2}} \exp\left[-\frac{T^2}{4}(\omega - \omega_0)^2\right]. \tag{4.5}$$

This is also Gaussian in form, centred about the carrier frequency ω_0 with a semi-width of

$$\delta = \omega - \omega_0 = \frac{2}{T}.$$

Thus the sharper the pulse, the greater the bandwidth (as measured by δ) and hence we should expect greater dispersion effects.

For example, the transmitted pulses at our Laboratory are approximately $T = 125$ μsec in semi-width, so that $\delta = 2{\cdot}5$ kc/s, or a total bandwidth of 5 kc/s. This is well within the 10 kc/s margin about the carrier which our computations of the previous section were based on.

The inverse (4.4) by the Fourier integral theorem is given by

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(\omega)\, e^{j\omega t}\, d\omega. \tag{4.6}$$

By (4.5) this becomes

$$f(t) = \frac{AT}{2\sqrt{\pi}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{4}T^2(\omega - \omega_0)^2 + j\,\omega t\right] d\omega. \tag{4.7}$$

If we denote by $X(\omega)$ the amplitude of the received wave to a unit incident c.w. of frequency ω, then we will have for the received pulse

$$F(t) = \frac{AT}{2\sqrt{\pi}} \int_{-\infty}^{\infty} X(\omega) \exp\left[-\frac{1}{4}T^2(\omega - \omega_0)^2 + j\,\omega t\right] d\omega$$

$$= \exp(j\,\omega_0 t)\, \frac{AT}{2\sqrt{\pi}} \int_{-\infty}^{\infty} X(\sigma) \exp\left[-\frac{1}{4}T^2\sigma^2 + j\,\sigma t\right] d\sigma, \tag{4.8}$$

where $\sigma = \omega - \omega_0$. The second form shows the carrier $\exp(j\,\omega_0 t)$ explicitly, modified by a certain form factor. To calculate $X(\omega)$ or $X(\sigma)$ from a discrete known set of

values $X(\omega_1)$, $X(\omega_2)$, ... we make use of the Hermite functions $\Phi_n(X)$ defined by

$$\Phi_n(x) = \exp(-\tfrac{1}{2}X^2)\, H_n(x),$$

where

$$H_0(x) = 1, H_1(x) = 2x, H_2(x) = 4x^2 - 2, \ldots, H_n(x) = (-1)^n \exp(x^2)\, d^n \exp \frac{(-x^2)}{dx^n}.$$

The Fourier transform operation is particularly simple with these functions, namely

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi_n(k\omega) \exp(j\omega t)\, d\omega = \left(\frac{j^n}{k}\right) \varphi_n\left(\frac{t}{k}\right) \tag{4.9}$$

(for this and other properties of Hermite functions, see reference [5]). If we now expand $X(\sigma)$ in a convergent Fourier-Hermite series of the form

$$X(\sigma) = \sum_0^{\infty} C_n H_n(\lambda) \qquad \lambda = \frac{T\sigma}{\sqrt{2}}$$

then

$$F(t) = \exp(j\omega_0 t) \frac{AT}{2\sqrt{\pi}} \sum_0^{\infty} C_n \int_{-\infty}^{\infty} \varphi_n(\lambda) \exp(j\sigma t)\, d\sigma$$

$$= \exp(j\omega_0 t)\, A \sum_0^{\infty} j^n C_n \varphi_n\left(\frac{t\sqrt{2}}{T}\right)$$

$$= A \exp\left(-\frac{t^2}{T^2} + j\omega_0 t\right) \sum_0^{\infty} j^n C_n H_n\left(\frac{t\sqrt{2}}{T}\right).$$

The terms in the infinite series in this expression represent successive orders of distortions of the transmitted pulse. Using the explicit values of H_n, this series becomes

$$C_0 + 2\sqrt{2}\, j C_1 \frac{t}{T} - 2C_2\left(4\frac{t^2}{T^2} - 1\right) - 4\sqrt{2}\, j C_3\left(4\frac{t^3}{T^3} - 3\frac{t}{T}\right) + \cdots.$$

Since the φ_n's form an orthonormal set, the coefficients are given by

$$C_n = \int_{-\infty}^{\infty} X(\sigma) H_n(\lambda) \exp\left(-\frac{1}{2}\lambda^2\right) \frac{d\lambda}{\sqrt{2^n n!\sqrt{\pi}}}.$$

In practical cases, however $X(\sigma)$ is known only for a discrete set of frequencies, in which case the C_n's can be evaluated more directly. If we suppose the power-series expansion $X(\sigma) = a_0 + a_1\sigma + a_2\sigma^2 + \cdots$ and use the relations

$$1 = H_0(\lambda) \qquad \sigma^2 = \frac{(\frac{1}{2}H_2(\lambda) + H_0(\lambda))}{T^2}$$

$$\sigma = \sqrt{2}\frac{H_1(\lambda)}{2T} \qquad \sigma^3 = \sqrt{2}\frac{!(\frac{1}{4}H_3(\lambda) + \frac{3}{2}H_1(\lambda))}{T^3}$$

we have, on rearranging terms,

$$X(\sigma) = \left(a_0 + \frac{a_2}{T^2} + \cdots\right) H_0(\lambda) + \left(a_1\frac{\sqrt{2}}{2T} + 3a_3\frac{\sqrt{2}}{2T^3} + \cdots\right) H_1(\lambda) +$$

$$+ \left(\frac{a_2}{2T^2} + \cdots\right) H_2(\lambda) + \left(a_3\frac{\sqrt{2}}{4T^3} + \cdots\right) H_3(\lambda) + \cdots.$$

If we neglect all coefficients a_n beyond $n = 2$, the pulse output becomes

$$A \exp\left(-\frac{t^2}{T^2} + j\omega_0 t\right)\left[\left(a_0 + \frac{2a_2}{T_2}\right) + 2j a_1 \frac{t}{T^2} - 4a_2\frac{t^2}{T^4}\right]. \tag{4.10}$$

VOL. 2 1952

To obtain the a_n's, suppose we know the three values $X(-\sigma_1)$, $X(0)$, $X(\sigma_1)$. Then we can use the interpolation formulas

$$a_0 = X(0), a_1 = \frac{[X(\sigma_1) - X(-\sigma_1)]}{2\sigma_1}$$

$$a_2 = \frac{[X(\sigma_1) + X(-\sigma_1) - 2X(0)]}{2\sigma_1^2}.$$

Let us apply these relations, for example, to the data given by (3.7). We have $\sigma_1 = 2\pi \times 10^4$ corresponding to 10 kc/s, and, for a unit input pulse exp $(-t^2/T^2 + j\omega_0 t)$ the output pulse through the coupling region was calculated from (4.10). The time varying terms turned out to be negligibly small, even for pulses as short as $T = 25$ μsec., However the constant term in the series contains the term $2a_2/T^2$ which is appreciable for short pulses. For example, we have, from the data of (3.9), the pulse amplitudes. Thus, shortening the pulse length below 150 μsec increases the amplitude transmitted through the layer.

Table 2. Pulse amplitudes as function of pulse lengths

T μsec	$a_0 + \frac{2a_2}{T^2}$
25	0·349
40	0·326
150	0·320

5. Dispersion in the Reflection Region

It is well-known that abnormal changes in group velocity with consequent effect on group height calculations occur near a height of reflection. An exact calculation would involve higher order terms in the expansion of the frequency as a function of the wave number. However, for the present problem a detailed wave-theoretical analysis is not required. We describe here a procedure which estimates this delay, and which is suggested by comparing the reflection process with a delay network in circuit theory. Fig. 2 shows how the indices of refraction $n_i = \mu_i - j\chi_i$ $(i = 1, 2)$ vary between X = −0·40 and X = −2·20. For our model assuming a scale height of 10 km this would correspond to the actual height interval 92–115 km. We consider this as the reflection region for the ordinary wave $(i = 1)$ since the values of μ_1 drop nearly to zero over this interval. If we form the extrapolated intercept of the μ_1-curves with the zero axis, as shown by the dotted lines of Fig. 2, this would determine a ray theoretical height of reflection. While this may deviate slightly from wave-theory heights, it is the height differences which we seek, and we have found that the method gives these differences fairly accurately. The χ_1-curves (absorption) show similar behaviour, and we may assume with sufficient accuracy that

Table 3. Phase heights vs incident frequency

f (kc)	h_ω (km)
140	106·3
150	107·6
160	109·0

$$\int_0^{h_\omega} \chi_1(\omega)\, dz = \text{const } 140\,\text{kc} \leq \frac{\omega}{2\pi} \leq 160\,\text{kc}$$

where h_ω is the (extrapolated) height of reflection defined above. The value of this integral is about 1 neper. (See reference [1].) The values of h_ω are given in Table 3. This corresponds to a path difference of 0·27 km per unit kc, a time difference of

$$\Delta t = \frac{0{\cdot}27}{3 \times 10^5} = 0{\cdot}9\,\mu\text{sec per unit kc}$$

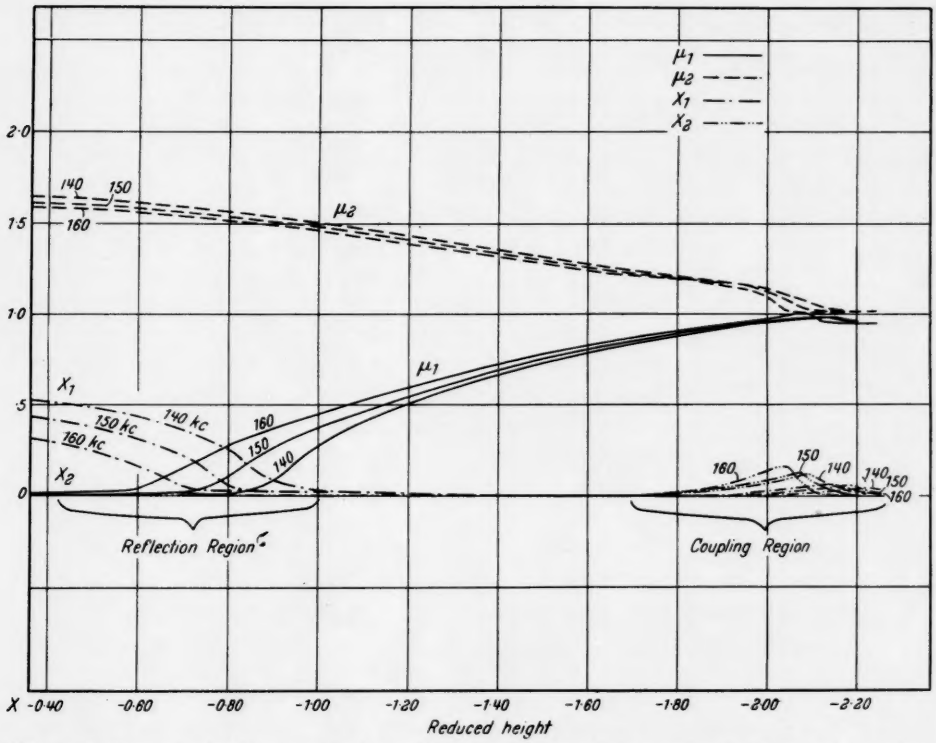


Fig. 2. Indices of refraction $\mu - j\chi$ for ordinary and extraordinary waves in E-reflection region for 0·55 mc/s critical frequency.

or a phase difference of

$$\varphi = \omega\,\Delta t = (150 \times 2\pi \times 10^3) \times 0{\cdot}9 \times 10^{-6} = 0{\cdot}85 \text{ rad. per kc}.$$

We may think of the situation in terms of a transmission line or filter with definite response characteristics. These are shown in Fig. 3. Our response function, as defined in (4.8) is then of the form

$$X(\omega) = Re^{-j\varphi} = Re^{-jk(\omega-\omega_0)}$$

that is, each frequency suffers a phase delay proportional to its deviation from the carrier frequency. Relation (4.8) then becomes

$$F(t) = \frac{ATR}{2\sqrt{\pi}}\, e^{j\omega_0 t} \int_{-\infty}^{\infty} \exp\left[-\frac{T^2}{4}\sigma^2 + j\,\sigma\,(t-k)\right] dk.$$

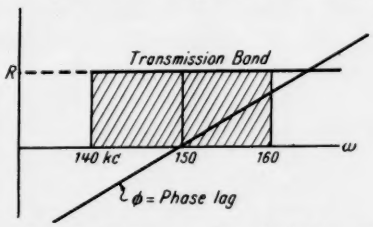


Fig. 3. Response characteristics of E-reflection region.

Putting $t-k$ for t gives

$$F(t) = RA\exp\left[\left(\frac{t-k}{T}\right)^2 + j\,\omega_0\,t\right]$$

$$= RA\exp\left[\left(\frac{t-k}{T}\right)^2 + j\,\omega_0\,(t-k)\right] \times e^{j\omega_0 k}.$$

VOL. 2 1952

This is a Gaussian pulse shifted through time $k = \frac{\Delta\varphi}{\Delta\omega}$, if we ignore the phase shifting factor $e^{j\omega_0 k}$. From the constants above we have

$$k = 150 \times 0{\cdot}9 \times 10^{-6} = 135 \times 10^{-6} \text{ or } 135\,\mu\text{sec}.$$

We have shown that our pulse suffers a time delay of 135 μsec at the reflection layer.

It thus appears that the reflection region may be said to behave as a delay line without distortion. This situation may have some bearing on the interpretation of the $h' - t$ records being obtained here at the Laboratory. These records show a variation of apparent reflection layer height with time of day. If we interpret this as due to change in critical frequency, we should seek confirmation on the records of the law of variation of this time delay with critical frequency. This has been calculated, using the approximate relation

Table 4. Time delay vs critical frequency

Critical frequency f_c *mc/s*	*Time delay* μsec
0·5	135
1·1	24
2·2	20
3·0	20
4·4	20

$$\varepsilon_1 \cong 1 - \frac{1}{\omega} \frac{\omega_c^2 ch\,(x)}{\omega\left(1 + \frac{4{\cdot}68 \times 10^{12}}{\omega_c^2 ch\,(x)} + 8{\cdot}90 \times 10^6\right)}$$

where $ch\,(x) = \frac{N(x)}{N_{\max}(x)}$ is the so-called "CHAPMAN" function. The results are (see Table 4). These are plotted in Fig. 4. It is of interest to note that we have a practically constant time-shift over most of the critical frequency range but this increases rather suddenly

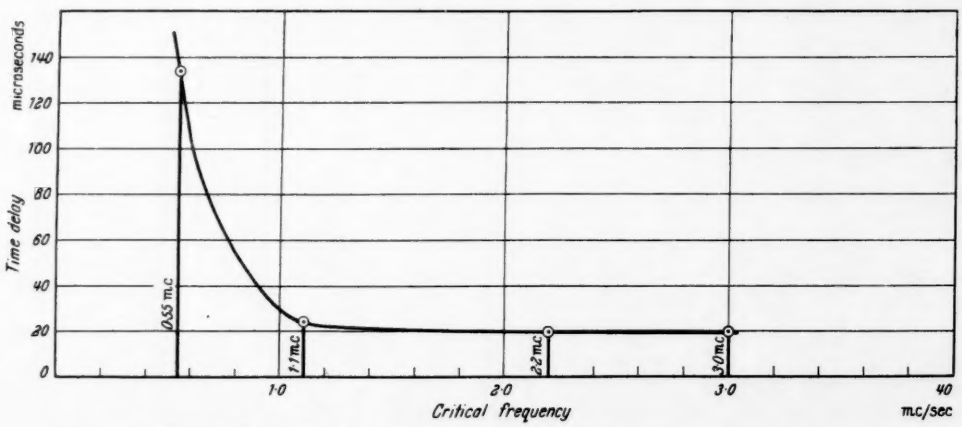


Fig. 4. Time delay for E-reflection as function of critical frequency.

more than five-fold at very-low criticals. It is likely that the critical frequencies drop down to this low value during the night-time hours, although its exact variation with time (during the night) is still uncertain. Some light can be shed on these questions by comparing group-height records against phase heights as obtained at 150 kc/s at the Laboratory (see Fig. 5). The night of November 25, 1951 was chosen because readable records of both types were available. The baseline for the phase height data was chosen both from the calculated theoretical values (Fig. 6), based on our present model, and by comparison with the monthly-

averaged daytime *E*-layer critical frequencies appearing in CRPL-F. The two agree closely. There is a consistent difference of about 14 km between the two from 2 a.m. until sunrise. If we assume the critical to stay constant at about 0.55 mc/s over this period, then our time-delay mechanism appears to fit the experimental data.

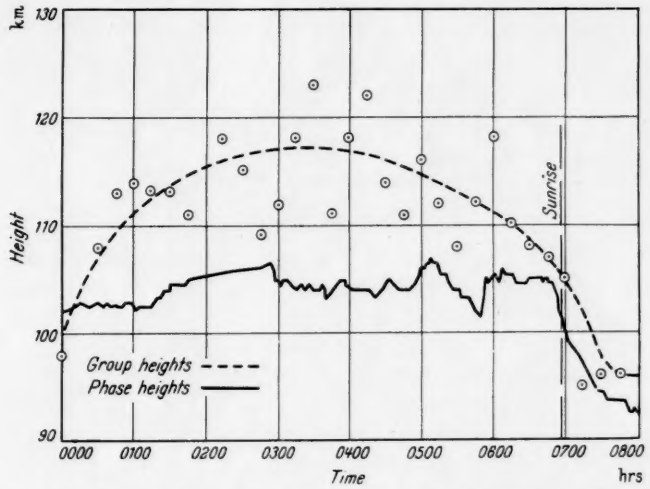


Fig. 5. Group and phase heights as a function of time 0000–0800 25 November 1951.

VOL. 2 1952

That the critical should behave as supposed is consistent from the physical viewpoint. For, at $f_c = 1$ mc/s, the number of electrons is $N_e = 1{\cdot}24 \times 10^4$ so

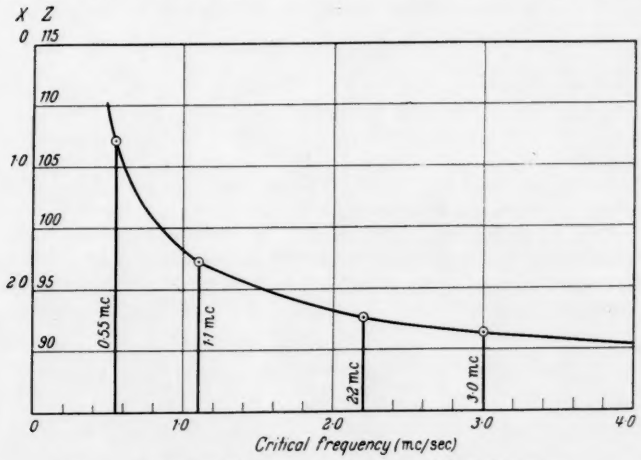


Fig. 6. Phase height as function of critical frequency.

that for slightly lower critical we have roughly $N_e \sim 10^4$. Disregarding the possible effect of sporadic ionization we may assume no ionization is taking place during the night so that we have

$$\frac{dN_e}{dt} = -\alpha N_e^2$$

where α is the effective recombination coefficient. If t be measured from the time when the critical first fell below 1 mc/s then the solution to the above equation is

$$N_e = \frac{1}{\alpha t + 10^{-4}}.$$

NICOLET suggests that $\alpha = 5 \times 10^{-9}$ is a safe maximum value to assign so that it requires

$$t = 2 \times 10^4 \text{ sec} = 6 \text{ hrs}.$$

before N_e drops by half or f_c by 30%. This time is of the order over which the effects shown in Fig. 5 take place.

CONCLUSIONS

While the Fourier integral is an exact tool for extending the magneto-ionic theory to pulse transmission, very often simpler approaches can be used, especially as distortions are small for pulses of the order of 100–200 μsec. Such distortion as may be present tends to concentrate around the pulse edges. It is inaccurate, therefore, to use these edges as time markers—the central maximum of the pulse should be used instead. The simpler mechanism of the delay line can be used to study E-layer reflection of pulses—here the phenomenon appears to be entirely one of phase shift with no amplitude distortion.

Thanks are due to ROBERT L. SCHRAG of this Laboratory who supplied much of the dispersion calculations for this paper, to Dr. A. H. WAYNICK, M. NICOLET, and others of this Laboratory for helpful advice and criticism, and especially to RUNE LINDQUIST and R. JONES for help in supplying and interpreting group and phase-height records.

REFERENCES

[1] GIBBONS, J. J. and NERTNEY, R. J.; "Wave Solutions, Including Coupling, of Ionospherically Reflected Long Radio Waves for a Particular E-region Model", to appear in J. Geophys. Res., September 1952. [2] GIBBONS, J. J. and NERTNEY, R. J.; "A Method for Obtaining the Solutions of Ionospherically Reflected Long Waves", J. Geophys. Res., 1951 **56** 355–371 (September). [3] RYDBECK, O. E. H.; "On the Propagation of Radio Waves", Trans. Chalmers. Inst. Tech., (Gothenburg) No. 34. [4] DAVIDS, NORMAN; "A Geometric Interpretation of the π-Wave and Coupling Factor in Ionospheric Long-Wave Theory", J. Geophys. Res., 1951 **56** 611 (Letter to the Editor). [5] COURANT, R. and HILBERT, D.; "Methoden der Mathematischen Physik, I." Springer, Berlin pp. 78–79. [6] RYDBECK, O. E. H.; "A Theoretical Survey of the Possibilities of Determining the Distribution of the Free Electrons in the upper Atmosphere." Trans. Chalmers. Inst. Tech., (Gothenburg), No. 3 (1942).

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 337 to 339. Pergamon Press Ltd., London

Electric charges from ice friction

J. Alan Chalmers, M.A., Ph.D., F.Inst.P.,
Senior Lecturer in Physics, Durham Colleges, University of Durham

(Received 20 May 1952)

ABSTRACT

Experiments are described showing that, when ice is rubbed, the large fragments carry a negative charge, in agreement with the theory of Simpson and Scrase [1] for thunderstorm electrification. The present experiments extend the results of Pearce and Currie [2] since they show charge separation even in the absence of air blasts.

VOL.
2
1952

1. Introduction

In 1937, Simpson and Scrase [1], from their observations with the alti-electrograph, found that the main separation of charge in a thundercloud occurs at temperatures below the freezing point, giving an upper positive and lower negative charge. They put forward the theory that the process of separation is concerned with friction between particles of ice, the larger fragments carrying a negative charge, while the corresponding positive charge is carried up in the rising air. This theory agreed with the fields found in blizzards, but lacked confirmation from laboratory experiments.

It is surprising to find that, in the 15 years since the theory was put forward, there has been no direct evidence that the simple friction of ice upon ice gives charges of the right sign. In fact, in one case the charge is of the wrong sign! Chapman [3] found that when snow struck snow in a dish, the dish acquired a positive and the air a negative charge.

The nearest investigations to the direct establishment of the frictional effect were those of Pearce and Currie [2], who eroded a block of snow with an air blast and found the larger eroded fragments to carry a negative charge, the air receiving a positive charge; they found the effect to be small until snow surfaces were fractured. Their measurements in blizzards gave similar results. While these results may be concerned with friction between particles of ice, the electrical effects might be explained merely by the action of the air blast on the snow.

Other measurements that have been reported are not of direct bearing. Stäger [4], [5] in two papers reported contradictory results. Schaefer [6] measured the charge produced on a metal surface by the impact of fast moving snow particles, but apparently did not measure the charges either on the fragmented snow or in the air. Measurements in which snow or ice grows by the impact of water droplets, *e.g.* Lueder [7], do not enter into the present problem.

The present work describes simple measurements which confirm that simple ice friction does give the charge separation postulated by Simpson and Scrase [1].

2. Experimental Details

Apparatus, which will be described in detail in another publication, had been set up for the measurement of rain charges and suitable conditions made it possible to investigate ice friction.

Snow, which was lying to a depth of a few inches, had formed an icy crust which made it particularly suitable for the investigations, in which it was necessary for fragmented ice to fall into the receiver after rubbing. Ice produced in a refrigerator does not fragment easily on rubbing and freshly fallen snow is too easily caught by a wind, but this icy snow was ideal for the purpose.

A handful of snow (50–100 g) was taken and rubbed together at a height of some 10 to 20 cm above the collector so that the fragments of snow fell into the collector; an appreciable breeze was blowing, so that any charge left in the air would be blown away over the collector.

The results were that on every occasion an amount of charge between -2 and -8×10^{-10} coulombs arrived at the collector; it is probable that the divergence between the separate measurements was due to differences in the quantities of snow used and in the amount of rubbing that had taken place.

Before the observations can be taken as giving definite proof of charging by ice friction, certain other possibilities must be eliminated, and some subsidiary experiments were carried out.

The charge might be concerned with the snow itself in some way, and not with the friction. For example, the snow might carry and retain the negative charge induced on the earth's surface by the positive field in the atmosphere. Accordingly, a handful of snow was lifted and dropped with as little friction as possible into the collector; the negative charge obtained in this way was only a few per cent of that obtained from a similar quantity of snow rubbed.

The effect might be caused by the impact of the snow on the metal of the collector. So the collector was coated on the inside with snow; the measurements showed no significant difference from those when the snow fell on the bare metal.

The charge might have been produced in some way by the hands on the ice; a handful of snow was taken and consolidated by the hands into a ball, with much more working by the hands than in the rubbing experiment. When the ball was carefully dropped into the collector, the charge obtained was very small.

The fragments of ice might get charged by induction by the earth's electric field in the same way as the drops in the water dropper. So a collector was used indoors and measurements taken. There was still a considerable negative charge from rubbing, though probably rather less than outside, perhaps because of the absence of wind to remove the positively charged air. The use of a different collector in this measurement eliminates any possible specific effect of the collector.

3. Discussion

The results establish that the friction of ice on ice gives rise to a negative charge on the larger particles of ice, of the order of 10^{-11} coulombs/g, under the conditions of the experiment, in the absence of any air blast.

This agrees with the fields found in blizzards and fits the ice-friction theory of Simpson and Scrase [1] in a qualitative way. The charge per g required to give thunderstorm phenomena would be very many times that found in the present measurements, but the conditions within a thunder cloud are so vastly different from those of the present work that it is quite impossible to derive quantitative conclusions and without these we cannot usefully discuss whether ice-friction can account for thunderstorm electricity.

However, one remark by PEARCE and CURRIE [2] requires comment. Since they found naturally falling snow flakes to be predominantly positively charged, they conclude that these cannot have obtained their charges by ice friction and so that ice friction cannot be the main feature of thunderstorm electrification. But SIMPSON [8] found that snow shows the "mirror image effect" even more markedly than rain, and it seems probable that the mirror image effect must be interpreted that the precipitation receives its charge, not in the cloud, but as a consequence of the field below. The author [9] has shown that WILSONs process is able to account for the mirror image effect for rain, and if, as suggested by the author [10], WILSONs process is also effective for snow, a similar account should be possible. Thus the charge on falling snow depends on the field existing and not on the processes at work in the cloud, and cannot give information as to whether ice-friction operates within clouds.

Acknowledgement—The author wishes to thank the Council of the Durham Colleges for the grant of a sabbatical year which has allowed the present work to be carried out without interruption from other duties.

REFERENCES

[1] SIMPSON, SIR G. C. and SCRASE, F. J.; Proc. Roy. Soc. A 1937 **161** 309. [2] PEARCE, D. C. and CURRIE, B. W.; Can. J. Res. 1949 **27** A 1. [3] CHAPMAN, S.; Report of Conference on Thunderstorm Electricity, p. 153. Chicago 1950. [4] STÄGER, A.; Ann. d. Phys. 1925 **76** 49. [5] STÄGER, A.; Ann. d. Phys. 1925 **77** 225. [6] SCHAEFER, V. J.; Trans. Amer. Geophys. Un. 1947 **28** 587. [7] LUEDER, H.; Z. Angew. Phys. 1951 **3** 247. [8] SIMPSON, SIR G. C.; Geophys. Mem. 84, 1949. [9] CHALMERS, J. A.; Q. J. Roy. Met. Soc. 1951 **77** 249. [10] CHALMERS, J. A.; Q. J. Roy. Met. Soc. 1947 **73** 324.

VOL.
2
1952

Journal of Atmospheric and Terrestrial Physics ,1952 ,Vol. 2 ,pp. 340 to 349. Pergamon Press Ltd., London

Über die Ursache der exzessiven Absorption in der Ionosphäre an Wintertagen

On the causes of Excessive Absorption in the Ionosphere on Winter Days

Walter Dieminger

Institut für Ionosphärenforschung in der Max-Planck-Gesellschaft, Lindau a. Harz

(Received 30 March 1952)

ABSTRACT

Excessive absorption is observed in the northern hemisphere on groups of days in winter. At the same time echoes are present at heights between 75 and 90 km on frequencies from 1·6 to 4 Mc/sec. These echoes show a regular diurnal variation of height with a minimum at noon. They are merging in a level of sporadic echoes at 95 km between sunset and sunrise. The height is independent of frequency. There is a close correlation between the height of the echoes and the absorption of short radio waves in the ionosphere. It appears that the echoes are produced by a partial reflection at a sharp boundary of an ionised region extending from *E*-layer down to heights between 70 and 90 km. The structure of the echoes points to inhomogeneous ionisation. Solar control is apparent but it is unlikely that the daily variation is due to changes of solar radiation. No conclusive explanation is available yet. The echo level at 95 km during night-time is attributed to meteoric dust ionisation.

In einer vorläufigen Mitteilung über die Beobachtung von Kurzwellenechos aus Höhen unter 100 km [1] wurde die Vermutung geäußert, daß ein Zusammenhang besteht zwischen dem Auftreten dieser Echos und der ungewöhnlich hohen Absorption, die elektrische Wellen in der Ionosphäre an bestimmten Tagen im Winter erleiden. Versuche, die in der Zwischenzeit angestellt wurden, haben diese Vermutung bestätigt. Die Ergebnisse seien im folgenden mitgeteilt.

1. Absorptionserscheinungen

Es ist seit langem bekannt, daß sich der zeitliche Gang der Absorption in der Ionosphäre nicht als eine einfache Funktion des Sonnenstandes ausdrücken läßt [2]. Das Gesetz, das aus der tageszeitlichen Änderung abgeleitet wurde [3], ergibt für den Winter einen zu kleinen Wert, d. h. die Absorption ist im Winter beträchtlich höher, als nach dem Sonnenstand zu erwarten ist. In der Praxis trägt man dem Rechnung, indem man für den Winter eine andere, empirisch gefundene Abhängigkeit ansetzt [4]. Woher die Diskrepanz kommt, war jedoch bislang völlig ungeklärt. Nun hat kürzlich Lange-Hesse [5] gezeigt, daß der jahreszeitliche Gang offenbar dadurch verwischt wird, daß bevorzugt im Winter „Ausreißer" auftreten, d. h. Tage, an denen die Dämpfung ein Vielfaches des erwarteten Wertes ist. Nach den Untersuchungen von Lange-Hesse treten die Tage mit hoher Dämpfung in nördlichen gemäßigten Breiten vor allem in den Monaten Dezember und Januar und zwar meist in Gruppen auf. Das sind die Monate, die auch in der Rawerschen Statistik [4] am weitesten herausfallen. Der Übergang zwischen Tagen mit hoher und niedriger Dämpfung kann dabei ganz abrupt erfolgen. Lange-Hesse hat weiterhin darauf hingewiesen, daß die Tage mit hoher Dämpfung mit *keiner bekannten Erscheinung* in der Ionosphäre zusammenhängen, wie z. B. Mögel-Dellinger-Effekte (S.I.D.), Ionosphärenstürme, sporadische *E*-Schicht. Bei unseren eigenen Feldstärkemessungen, die seit 3 Jahren laufend durchgeführt werden, hat das Auftreten dieser exzessiven Dämpfung in den Wintermonaten regelmäßig zu einer — natürlich erfolglosen — Fehlersuche in den Feldstärkemeßgeräten geführt. Auch auf den Echolotungsaufnahmen mit veränderlicher Frequenz war die erhöhte Dämpfung an der geringen Intensität der Echos deutlich zu erkennen.

2. Echoerscheinungen

Bereits im Jahre 1949 war aufgefallen, daß an manchen Tagen im Winter Echos unterhalb der *E*-Schicht auftreten. Abb. 1 zeigt einen Ausschnitt aus einer derartigen Aufnahme, bei der zunächst (linke Hälfte des Bildes) mit fester Frequenz (1,52 MHz) und dann (rechte Hälfte) mit veränderlicher Frequenz registriert wurde. Man erkennt deutlich unterhalb der *E*-Reflexionen, die bei etwa 1,7 MHz einsetzen, eine weitere Echospur. Sie ist bis etwa 3,5 MHz zu verfolgen, ihre Höhe ist unabhängig von der Frequenz etwas unter 100 km. Ähnliche Bilder traten auch im Winter 1950/51 auf. Die technischen Eigenschaften der Lotungsanlage gestatteten seinerzeit keine exakte Höhenbestimmung der tiefen Echos. Es konnte nur festgestellt werden, daß sie nicht identisch waren mit der sporadischen *E*-Schicht.

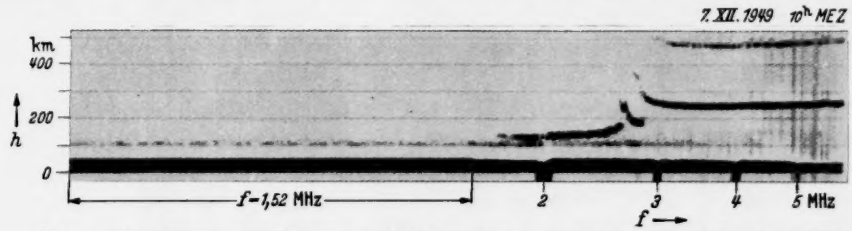


Abb. 1. Echolotungsaufnahme mit variabler Frequenz. Lindau, 7. 12. 1949. 10^{00} h MEZ.

VOL. 2 1952

Im Winter 1951/52 gelang es dann mit Hilfe einer Registrieranlage mit verbesserter Höhenauflösung, die charakteristischen Eigenschaften der Echos genauer zu studieren. Der zeitliche Ablauf läßt sich folgendermaßen beschreiben: Beobachtet man auf einer festen Frequenz von 2,0 MHz, so erhält man während der Nacht diffuse Echoerscheinungen von etwa 90 km aufwärts. Auf diese Echos soll am Schluß noch eingegangen werden. Kurz nach Sonnenaufgang löst sich hieraus ein Echoniveau, ab, das im Laufe einer Stunde um 10—20 km absinkt, im allgemeinen um Mittag seinen tiefsten Wert erreicht und gegen Sonnenuntergang wieder im 90-km-Niveau verschwindet (*D*-Echos). Derartige Registrierungen zeigen die Abb. 2, 3 und 4*. Das gleichzeitige Verhalten der *E*-Echos ist von Tag zu Tag verschieden. In manchen Fällen bleiben die *E*-Reflexionen fast ungeschwächt bestehen (Abb. 4), in anderen setzen sie in den Mittagsstunden vollständig aus (Abb. 2).

Weiteren Aufschluß geben Aufnahmen mit veränderlicher Frequenz (Abb. 5). Auch hier erkennt man, daß um Sonnenaufgang und -untergang (7^{40} h und 17^{20} h) neben den üblichen Echos aus der *E*- und *F*-Schicht ein verwaschenes Niveau bei etwa 95 km vorhanden ist. Um 9^{10} h sind bereits deutliche Echos aus rund 85 km Höhe vorhanden. Gegen Mittag (12^{30} h) ist die Erscheinung am deutlichsten und die Echohöhe am kleinsten. Um 15^{50} h hat die Intensität bereits wieder erheblich nachgelassen. Das Auffälligste ist, daß die Echos über einen breiten Frequenzbereich mit konstanter Höhe auftreten. Besonders deutlich ist das bei der Mittagsaufnahme, in der die *D*-Echos von 1,6 MHz (unterhalb sind Lotungen wegen des Rundfunkbetriebes nicht möglich) bis 4,5 MHz vorhanden sind. Ein kaum merklicher Anstieg der Höhe gegen hohe Frequenzen zu ist zweifellos durch die Abnahme der Echoamplitude vorgetäuscht.

* Man erkennt die Höhenänderung besonders deutlich, wenn man ein Lineal an die Unterkante der *D*-Echos anlegt.

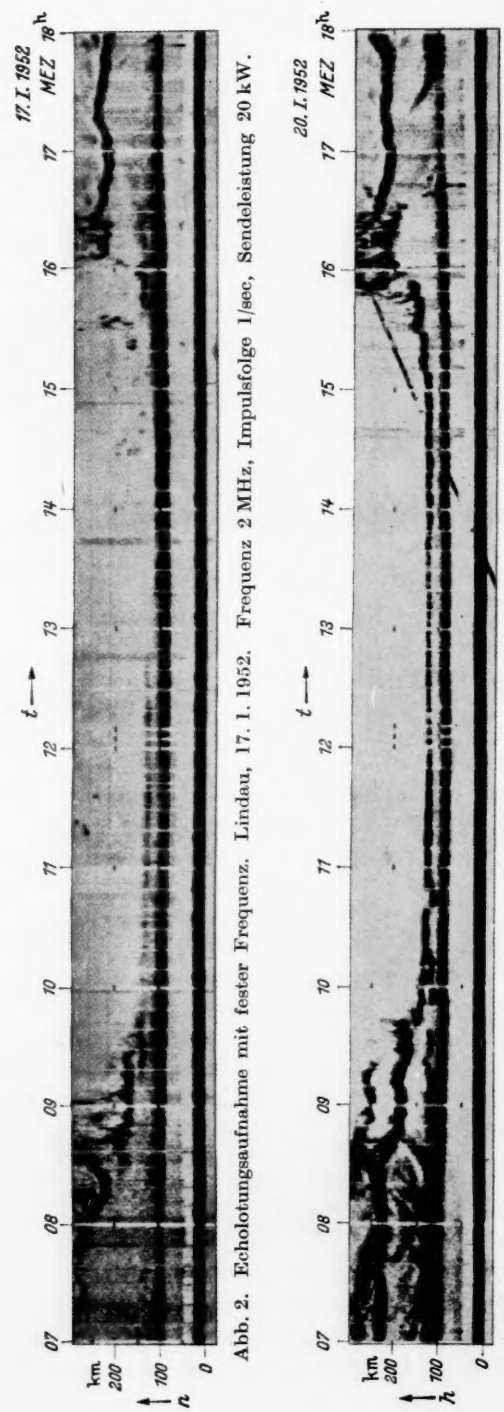


Abb. 2. Echolotungsaufnahme mit fester Frequenz. Lindau, 17. 1. 1952. Frequenz 2 MHz, Impulsfolge 1/sec, Sendeleistung 20 kW.

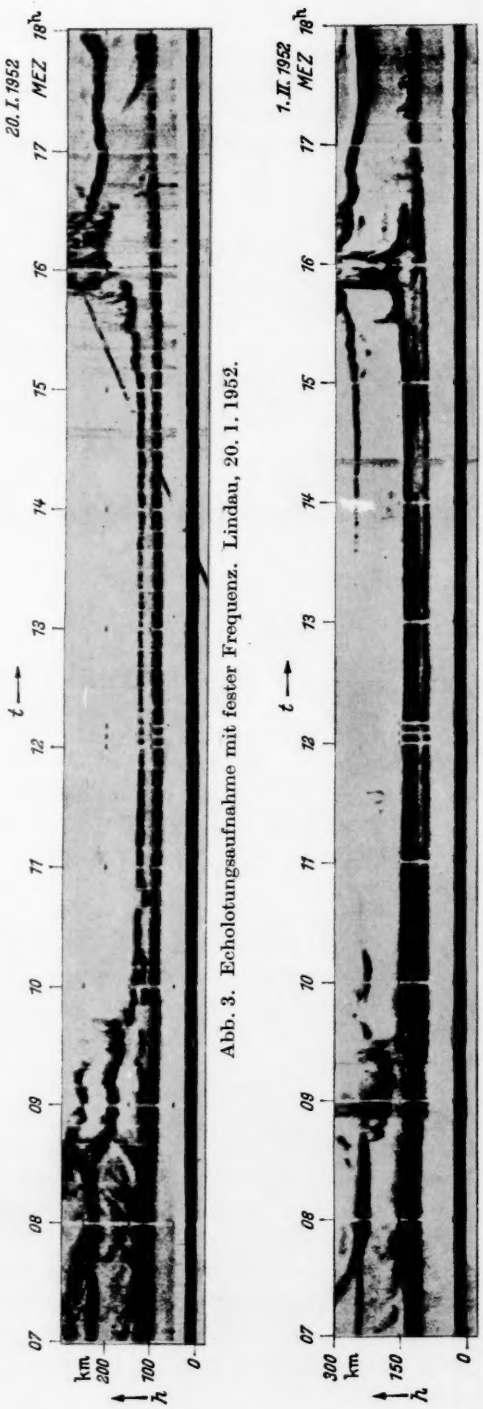


Abb. 3. Echolotungsaufnahme mit fester Frequenz. Lindau, 20. 1. 1952.

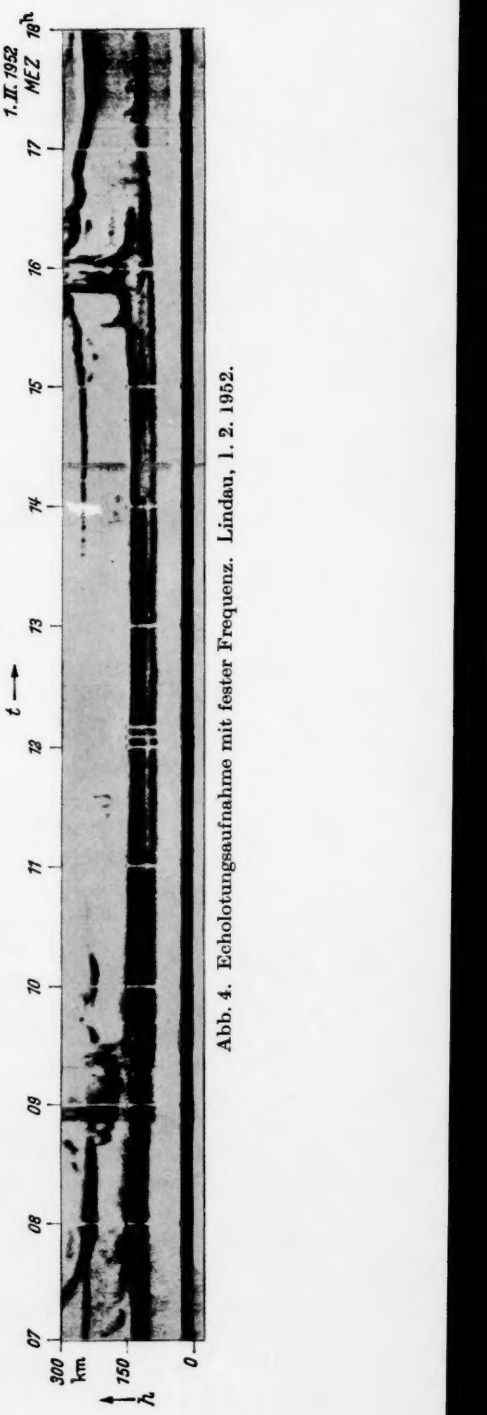


Abb. 4. Echolotungsaufnahme mit fester Frequenz. Lindau, 1. 2. 1952.

Ein Bild von der Veränderlichkeit der minimalen Höhe der *D*-Echos von Tag zu Tag gibt die Zusammenstellung in Abb. 6. Dort sind die Echohöhen für die Zeit vom 12. 1.—3. 2. 1952 in Abhängigkeit von der Tageszeit aufgetragen. Es ist zu

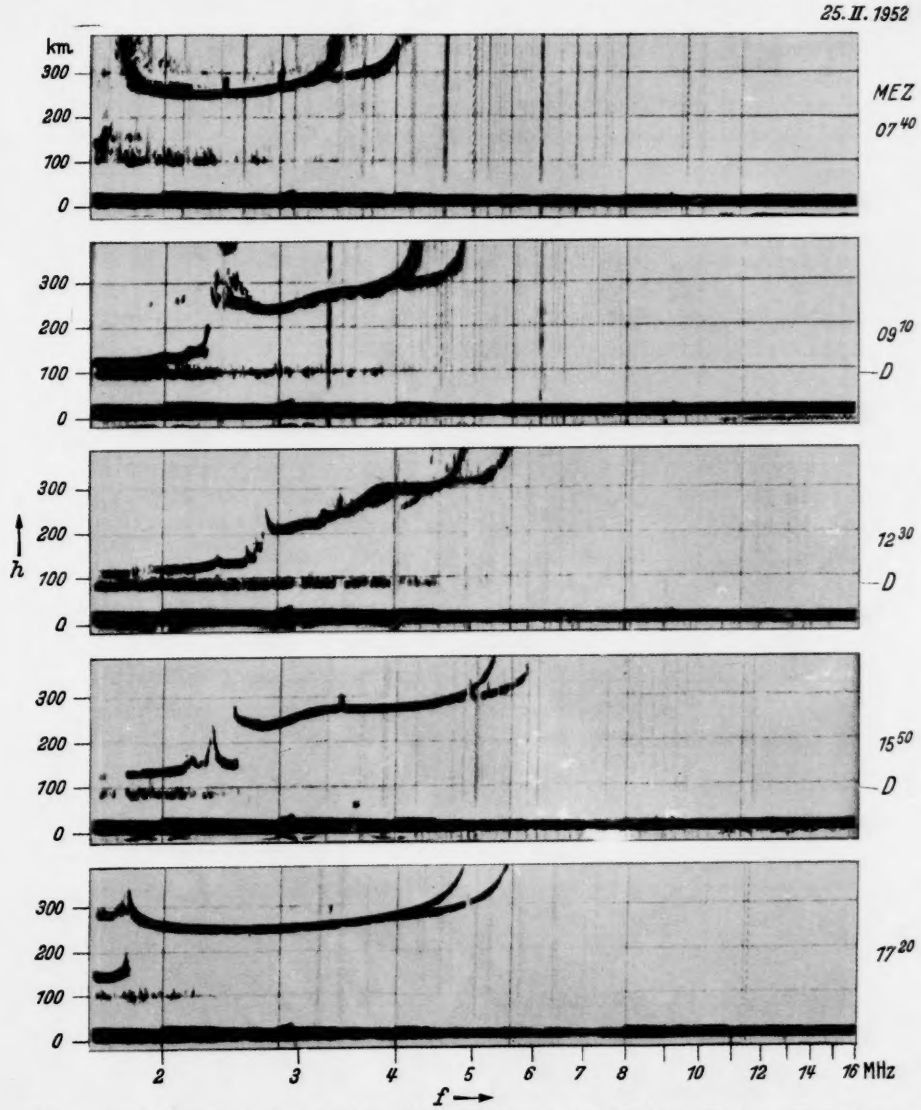


Abb. 5. Echolotungsaufnahmen mit variabler Frequenz (Ausschnitte). Lindau, 25. 2. 1952. Impulsfolge 50/sec, Sendeleistung 10 kW.

erkennen, daß der Tagesgang und der Tiefstwert, der im Laufe des Tages erreicht wird, individuell verschieden sind. Tage mit besonders tiefen Echos sind z. B. der 22., 24. und 26. 1. 1952, während am 3. 2. 1952 tagsüber überhaupt keine Echos unter 100 km auftraten.

VOL. 2 1952

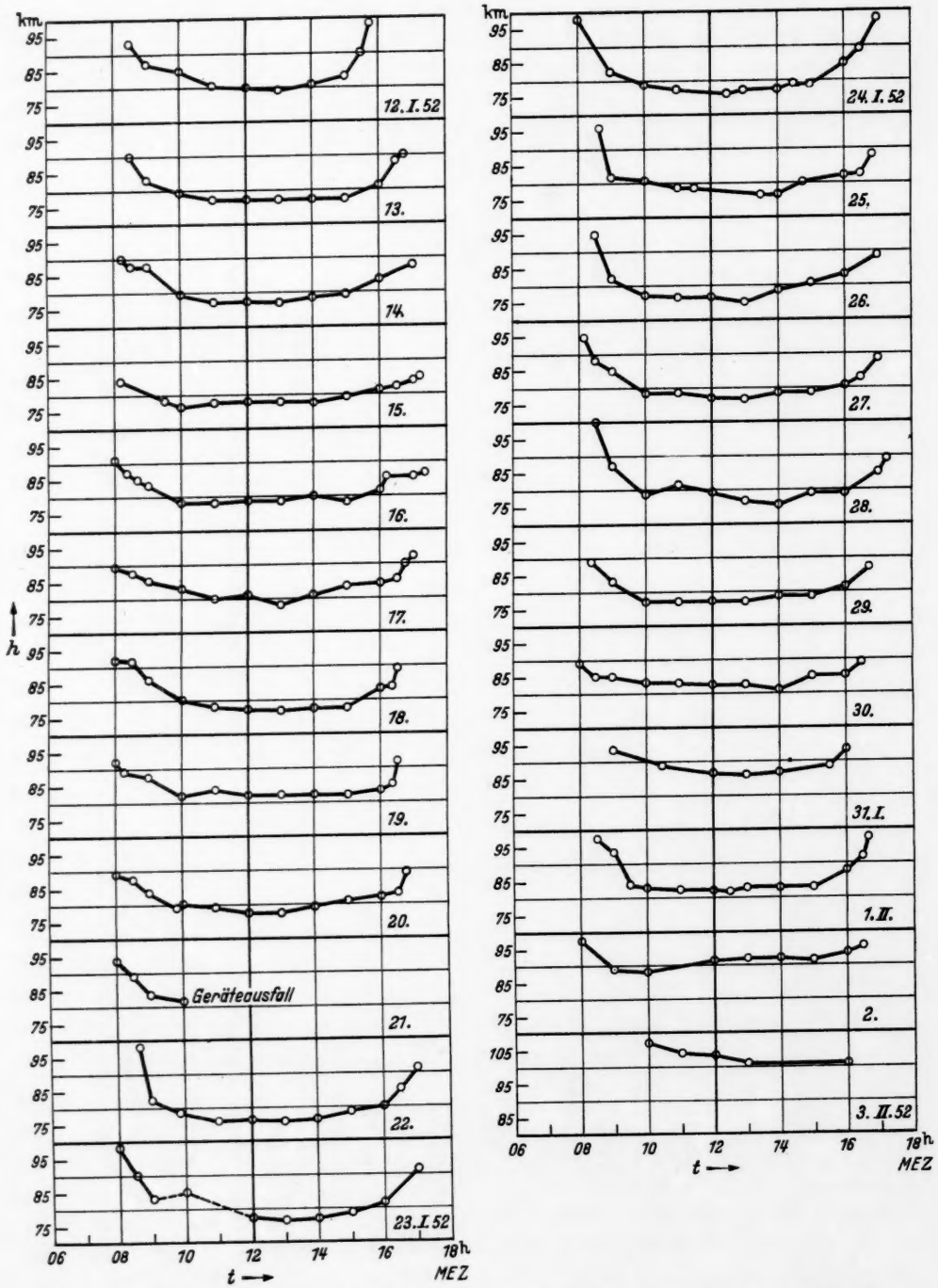


Abb. 6. Echohöhen in Abhängigkeit von der Tageszeit für den Zeitraum 12. 1.—3. 2. 1952.

3. Zusammenhang zwischen tiefen Echos und Absorption

Ein flüchtiger Vergleich von Echohöhen und Feldstärkeregistrierungen zeigte einen deutlichen Zusammenhang beider Phänomene. An Tagen mit niedrigen Echohöhen herrschten geringe Feldstärken, d. h. hohe Absorption, an Tagen ohne tiefe Echos dagegen hohe Feldstärken, d. h. geringe Absorption. Da jedoch Feldstärkeregistrierungen wegen der Vielzahl der enthaltenen Parameter für quantitative Untersuchungen wenig geeignet sind, wurde nach bekannten Vorbildern ein Gerät improvisiert, das gestattet, aus den Echoamplituden die Gesamtabsorption in der Ionosphäre zu bestimmen.

Es gelang gerade, einen Übergang von starker zu geringer Absorption zu erfassen. Für die wenigen Tage, an denen gleichzeitig Höhen- *und* Absorptionsmessungen durchgeführt werden konnten, ergibt sich nun ein überraschend enger Zusammenhang zwischen beiden Erscheinungen. Das Ergebnis ist in Abb. 7 dargestellt. Es wurde für jeden Beobachtungstermin am 29., 30. und 31. 1. 1952, für den gleichzeitig beide Meßwerte vorliegen, die Echohöhe als Funktion der Absorption eingetragen. Es kann natürlich in Anbetracht der wenigen Meßpunkte nicht behauptet werden, daß der Zusammenhang linear ist. Auch gilt das Ergebnis zunächst nur für den untersuchten Zeitraum. Für diesen ist aber der Zusammenhang ganz eindeutig.

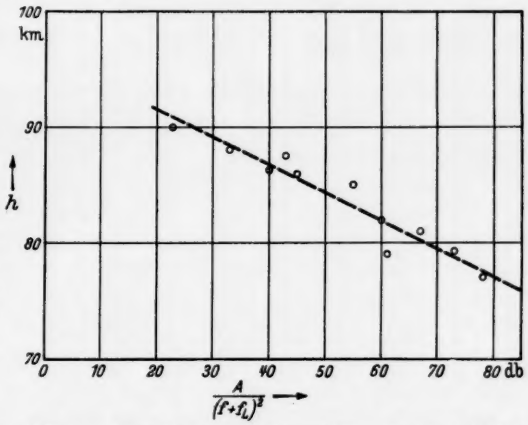


Abb. 7. Absorption in der Ionosphäre für $f = 2{,}0$ MHz als Funktion der minimalen Echohöhe für den Zeitraum 29. 1.—31. 1. 1952.

4. Versuch einer Deutung

Für eine Reflexion in der Ionosphäre bei senkrechtem Einfall gibt es 2 Grenzfälle: 1. Totalreflexion an der Stelle $n = 0$ (n Brechzahl). 2. Partielle Reflexion an einer Stelle, an der sich n auf der Strecke einer Wellenlänge merklich ändert *. Im zweiten Fall dringt ein Teil der Energie über die Reflexionsstelle hinaus in die Ionosphäre ein und kann unter Umständen an einer anderen Stelle total reflektiert werden. Um den ersten Fall kann es sich hier nicht handeln, denn gleichzeitig mit den *D*-Echos werden Echos aus den höherliegenden normalen Schichten der Ionosphäre beobachtet. Man könnte einwenden, daß es sich bei den *D*-Echos um Reflexionen der außerordentlichen Komponente handelt, während die gleichzeitigen Echos aus der *E*- oder *F*-Schicht der ordentlichen Komponente zuzuordnen sind. Dieser Einwand ist nicht stichhaltig wegen des bekannten Zusammenhangs zwischen den Reflexionsbedingungen der a. o. und o. Komponente. Nehmen wir einmal an, es handele sich in der Abb. 5 (Aufnahme von 12^{30} h) bei den *E*-Echos um die total reflektierte o. Komponente, bei den *D*-Echos dagegen um die total reflektierte a. o. Komponente. Da bei 1,6 MHz offenbar die o. Komponente die tiefe Schicht

* Auf den Übergang zwischen beiden Fällen, der z. B. von Rawer [6] ausführlich behandelt wurde, soll hier nicht eingegangen werden, da es für die folgende Argumentation ohne Einfluß ist.

bereits durchdrungen hat, müßte die a. o. Komponente das gleiche spätestens bei 1,6 + 0,7 MHz = 2,3 MHz tun. Tatsächlich sind aber die tiefen Echos bis zu 4,5 MHz vorhanden. Ferner hat BECKER [7] gezeigt, daß im Frequenzbereich zwischen der Gyrofrequenz und etwa 2,5 MHz bei Tage in der *E*-Schicht wegen der hohen Stoßzahl die Brechzahl für die a. o. Komponente gar nicht mehr Null wird. Das gilt in erhöhtem Maße für Gebiete, die tiefer als die *E*-Schicht liegen. Eine Totalreflexion an der Stelle $n = 0$ kann demnach für die a. o. Komponente unter diesen Umständen nicht auftreten. Damit bleibt nur die Annahme einer partiellen Reflexion an einer Sprungstelle der Brechzahl übrig. Mit dieser Annahme ist das Ergebnis der Aufnahmen mit veränderlicher Frequenz gut verträglich. Bei einer Sprungstelle erfolgt die Reflexion für alle Frequenzen in der gleichen Höhe. Die reflektierte Energie nimmt umgekehrt proportional der Frequenz ab, da der Unterschied in der Brechzahl mit zunehmender Frequenz immer kleiner wird.

Die Größe der erforderlichen Brechzahländerung läßt sich aus dem Reflexionskoeffizienten ϱ abschätzen, der seinerseits aus der Amplitude der tiefen Echos bestimmt werden kann. ϱ liegt nach unseren Messungen in der Größenordnung von 10^{-3}. Dies erfordert für n eine Änderung der gleichen Größenordnung. Bei Frequenzen um 3 MHz bedeutet das, daß sich die Elektronenkonzentration mehr oder weniger sprunghaft um etwa 1000 cm^{-3} ändern muß. Das kann z. B. an der Unterkante eines ionisierten Gebietes auftreten, das von dem Reflexionsniveau der tiefen Echos bis zur *E*-Schicht reicht.

Daß dieses Gebiet bei Tage bereits unter normalen Umständen ionisiert ist, geht aus der Reflexion der Längstwellen eindeutig hervor [8]. Die aus den Langwellenbeobachtungen abgeleiteten Elektronenkonzentrationen liegen bei $2 \cdot 10^2 \div 2{,}5 \cdot 10^3$ cm^{-3} in 70 km und bei $1{,}5 \cdot 10^4$ cm^{-3} in 90 km Höhe für die Mittagsstunden [9]. Die Ansichten über die Höhenverteilung sind geteilt. Manche Autoren [10] nehmen eine scharfe untere Grenze an, andere [11] rechnen mit einer CHAPMAN-Verteilung. Aus der Tatsache, daß an der Mehrzahl der Tage keine *D*-Echos im Kurzwellenbereich vorhanden sind, kann jedenfalls gefolgert werden, daß die Unterkante normalerweise nicht so scharf ist, daß sich die Elektronenkonzentration auf eine Strecke von 100 m ($\approx$ eine Wellenlänge) um 10^3 El/cm^3 ändert. Für die Erklärung der *tiefen Echos* würde die Annahme ausreichen, daß unter Umständen, und zwar bevorzugt im Winter, an einer Stelle der Gradient der Elektronenkonzentration diesen Wert erreicht. Es ist interessant, daß MITRA [12] auf rein theoretischem Wege ebenfalls zu einem beachtlichen Knick im Höhenverlauf der Elektronenkonzentration kommt, wenn er den Absorptionsquerschnitt des ionisierten Gases (O_2) zu $5 \cdot 10^{-22}$ cm^2 annimmt, daß diese Unstetigkeit aber für einen Absorptionsquerschnitt von 10^{-20} cm^2 verschwindet. Für die Erklärung der gleichzeitigen *hohen Absorption* genügt diese Annahme nicht. Hierzu ist erforderlich, daß in einem nicht zu schmalen Gebiet zwischen dem Reflexionsniveau der *D*-Echos und der *E*-Schicht die Elektronenkonzentration und damit die Dämpfung größer ist als normalerweise. Ein Modell, das den Beobachtungstatsachen gerecht würde, wäre demnach folgendes: Ein ionisiertes Gebiet, an dessen unterer Begrenzung ein ausreichender Gradient der Elektronenkonzentration vorhanden ist, und das mehr oder weniger tief unter die *E*-Schicht hinunterreicht. Wir sind uns darüber klar, daß diese Vorstellung rein qualitativ ist. Quantitative Untersuchungen sollen in weiteren Arbeiten erfolgen.

Über die Struktur des Ionisierungsgebietes läßt sich vorläufig nur sagen, daß die Unterseite offensichtlich keine glatte Fläche ist. Das läßt sich aus der zeitlichen Inkonstanz der tiefen Echos ableiten (Abb. 8). Ob das Ein- und Aussetzen der Reflexion durch Interferenz von Wellen verschiedener Laufzeit oder durch rasche Änderung des Reflexionsvermögens verursacht ist, läßt sich noch nicht entscheiden. Beides würde auf Inhomogenität der Ionisierung hindeuten*.

Ganz offen ist noch die Frage, wie diese Ionisierung entsteht und warum sie nur an bestimmten Tagen auftritt. Der tageszeitliche Gang der minimalen Höhe

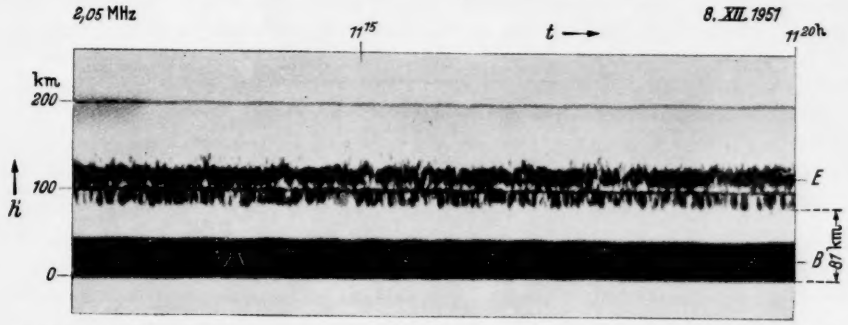


Abb. 8. Echolotungsaufnahme auf fester Frequenz mit schnellem Filmvorschub. Lindau, 8. 12. 1951. Frequenz 2,05 MHz, Impulsfolge 50/sec, Sendeleistung 20 kW; *B* Bodenwelle; *E* *E*-Reflexion.

beweist, daß eine solare Kontrolle vorliegt. Es ist aber unwahrscheinlich, daß die Änderungen von Tag zu Tag durch Schwankungen der solaren Strahlung, seien es Wellen oder Corpuskeln, bedingt ist. Gegen solare Wellenstrahlung spricht folgendes indirekte Argument: Die exzessive Dämpfung, die doch offenbar mit den tiefen Echos ursächlich zusammenhängt, tritt an verschiedenen Stellen der Erde *nicht* gleichzeitig auf. Das geht aus einem Vergleich der Dämpfungsmessung hervor, die seit mehreren Jahren in Slough (England), Singapore (Malaya) und auf den Falkland-Inseln durchgeführt werden**. Dies müßte aber der Fall sein, wenn solare Wellenstrahlung die dominierende Ursache wäre. Gegen solare Corpuskeln spricht der Umstand, daß keinerlei Zusammenhang gefunden werden konnte mit den Erscheinungen, die ziemlich sicher durch Corpuskeln hervorgerufen werden, nämlich erdmagnetische Störungen und Ionosphärenstürme. Es ist demnach die andere Alternative ernstlich in Erwägung zu ziehen, daß sich der Aufbau der Atmosphäre von Tag zu Tag so ändert, daß sich an manchen Tagen die beschriebene Ionisierungsverteilung in Höhen zwischen 75 und 100 km ausbildet, an anderen Tagen dagegen nicht. Die Konzentration der Erscheinung auf die Wintermonate kann vielleicht

* *Anmerkung bei der Korrektur:* Eine inhomogene Dämpfungsschicht würde auch erklären, warum die Abhängigkeit der Absorption vom Einfallswinkel nicht dem Martynschen Theorem gehorcht [21]. Für diesen Hinweis bin ich Sir Edward Appleton zu Dank verpflichtet. Inhomogenitäten in der dämpfenden Schicht könnten auch die Ursache dafür sein, daß die Methode von Bibl und Rawer zur Trennung des Absorptionsanteiles der *D*- und *E*-Schicht [22] in den Wintermonaten nicht zum Ziele führt, da bei dieser Methode mit homogener Verteilung der Elektronen gerechnet wird.

** Für die Überlassung dieser Werte bin ich dem *Department of Scientific and Industrial Research*, London, zu besonderem Dank verpflichtet.

VOL. 2 1952

einen Fingerzeig geben, in welcher Richtung zu suchen ist*. Genauere Aussagen über diese Zusammenhänge lassen sich erst machen, wenn Beobachtungen über die räumliche Ausdehnung des Phänomens vorliegen.

5. Vergleich mit bekannten Beobachtungen

Daß die hohe Atmosphäre herab bis zu 70 km ionisiert ist, geht — wie schon erwähnt — eindeutig aus der Reflexion von sehr langen Wellen hervor [8]. Es ergeben sich dabei ähnliche Höhen wie sie für die tiefen Echos gemessen wurden. Auch der Tagesgang des Reflexionsniveaus ist in ähnlicher Weise vorhanden. Im übrigen ist zu bedenken, daß die Reflexionsbedingungen für die beiden Wellengebiete durchaus verschieden sind. Eine Elektronenkonzentration bzw. ein Gradient der Elektronenkonzentration, der zur Reflexion der Längstwellen ausreicht, ist für Kurzwellen vollständig unwirksam. Immerhin wird es sich lohnen, das Verhalten der Längstwellen an Tagen mit tiefen Echos zu studieren.

Auch über niedrige Reflexionen bei Kurzwellen finden sich in der Literatur einige Angaben. Es ist nicht zu bezweifeln, daß diese Echos auch schon an anderen Stellen beobachtet wurden. So sind wahrscheinlich die Streuechos, die Eckersley [13] häufig in Höhen um 75 km von Impulssendern großer Leistung beobachtet hat, mit den hier beschriebenen identisch. Echos aus Höhen um 60 km sind in tropischen Gegenden von S. K. Mitra und Syam [14], sowie später von Ellyett [15] beobachtet worden. Letzterer gibt in seiner Arbeit eine Zusammenstellung älterer Beobachtungen; die angegebenen Reflexionshöhen sind jedoch 10—15 km niedriger als sie hier beobachtet wurden. Es ist daher zweifelhaft, ob es sich um das gleiche Phänomen handelt. Im übrigen schreibt A. P. Mitra [9] über diese Beobachtungen: „The behavior of the sporadic *D*-echoes are yet ill-understood, and, in many cases, there is considerable divergence of evidence." Dagegen haben Naismith und Bramley [16] bei ihren eingehenden Untersuchungen an Loran-Anlagen auf 2 MHz offensichtlich das gleiche Reflexionsniveau um 75 km beobachtet, das sich aus unseren Zenitmessungen ergibt. Auch Stranz [17] berichtet über Echos aus Höhen zwischen 85—100 km auf Frequenzen zwischen 3 und 5,8 MHz, doch handelt es sich dabei um Mögel-Dellinger-Effekte (S.I.D.). Über tiefe Echos, die in Cambridge im Winter beobachtet wurden, berichtet Whale [18]. Seine Angaben sind aber so knapp, daß man nicht entscheiden kann, ob es sich um die gleiche Erscheinung handelt. Über den tageszeitlichen Gang der tiefen Kurzwellenechos und den Zusammenhang mit der exzessiven Dämpfung im Winter ist bisher noch nichts bekanntgeworden. Daß sie von uns so eindeutig beobachtet werden konnten, liegt offenbar daran, daß die Empfindlichkeit unserer Echolotungsapparatur rund eine Zehnerpotenz größer ist, als die der üblichen Routinegeräte.

6. Nächtliche Echos aus Höhen unter 100 km

Abschließend sei kurz auf die — bereits eingangs erwähnten — streuenden Echos eingegangen, aus denen sich nach Sonnenaufgang die tiefen Echos herauslösen und in denen sie bei Sonnenuntergang wieder verschwinden. Folgende Eigenschaften deuten darauf hin, daß es dabei sich um die ionisierten Spuren von Meteorstaub

* *Anmerkung bei der Korrektur:* In diesem Zusammenhang interessiert die Mitteilung, daß im Winter 1951/52 an verschiedenen Tagen überraschend große Temperaturzunahmen (bis zu 40°) in der Hochstratosphäre gemessen wurden [23].

handelt: 1. Die Lebensdauer der individuellen Echos liegt in der Größenordnung einer Sekunde. 2. Die Unterkante der Echos fällt mit der unteren Grenze des relativ schmalen Bereiches [19] zusammen, in dem die ionisierten Spuren größerer Meteore festgestellt wurden. 3. Die Echos weisen eine statistische Häufung in den frühen Morgenstunden auf, also zu der Zeit, zu der die Erde dank ihrer Eigenbewegung am meisten kosmische Materie aufsammelt. Die Vermutung, daß es sich bei den nächtlichen Streuechos um Spuren von Meteorstaub handelt, wurde bereits früher vom Verfasser geäußert [20] und neuerdings von NAISMITH [16] bestätigt. In diesem

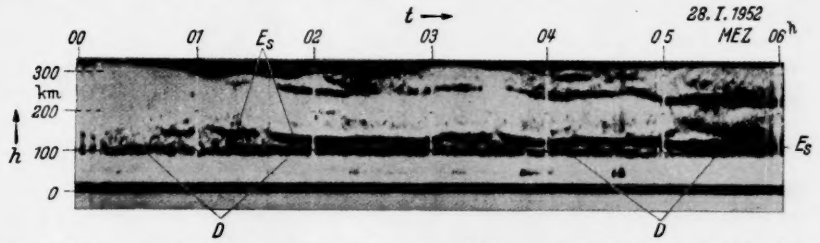


Abb. 9. Echolotungsaufnahme auf fester Frequenz. Lindau, 28. 1. 1952. Technische Daten wie Abb. 2. E_s sporadische E-Reflectionen, D Echos aus $\approx$ 95 km.

Zusammenhang interessiert noch, daß sich die sporadische E-Schicht in ihrer Struktur deutlich von diesen Echos unterscheidet (Abb. 9). Es handelt sich demnach wahrscheinlich um verschiedene Erscheinungen. Es ist aber durchaus möglich, daß bei Registrierungen mit geringer Höhenauflösung beide Phänomene miteinander vermischt werden. Insofern ist die Vermutung, daß die sporadische E-Schicht zum Teil auf Meteorstaub zurückzuführen ist, nicht unrichtig.

Meinen Mitarbeitern bin ich für die tatkräftige Hilfe bei diesen Untersuchungen zu besonderem Dank verpflichtet: Herr A. E. HOFFMANN-HEYDEN hat die Echolotungsapparatur auf ihren hohen Stand gebracht, Herr H. G. MÖLLER die Geräte für die Dämpfungsmessung improvisiert und Herr W. BECKER mich bei der Diskussion der Erscheinungen wesentlich unterstützt.

LITERATUR

[1] DIEMINGER, W. und HOFFMANN-HEYDEN, A. E.; Naturwiss. 1952 **39** 84. [2] APPLETON, E. V.; Proc. Roy. Soc. A 1937 **162** 451. [3] BEST, J. E. und RATCLIFFE, J. A.; Proc. Phys. Soc. 1938 **50** 233. [4] RAWER, K.; A. E. Ü. 1951 **5** 154. [5] LANGE-HESSE, G.; Tagung der Arbeitsgemeinschaft Ionosphäre Kleinheubach, August 1951. [6] RAWER, K.; Hochfr. u. Elektroak. 1939 **53** 150. [7] BECKER, W.; J. Atmosph. Terr. Phys. 1950 **1** 73. [8] BRACEWELL, R. N., BUDDEN, K. G., RATCLIFFE, J. A., STRAKER, T. W. und WEEKES, K.; Proc. Inst. Elect. Engrs. 1951 **98** (III) 221. [9] MITRA, A. P.; J. Geophys. Res. 1951 **56** 373. [10] BREMMER, H.; Terr. Radio Waves, Elsevier Publishing Co., Inc., New York 1949. [11] RYD-BECK, O. E. H.; Trans. Chalmers Univ., No. 34, 1944. [12] PFISTER, W.; Proc. of Conf. on Ionospheric Physics, State College, Penn., 1949. [13] ECKERSLEY, T. L.; J. Inst. Elect. Engrs 1932 **67** 992; 1937 **71** 405. [14] MITRA, S. K. und SYAM, P.; Nature Lond. 1935 **135** 953. [15] ELLYETT, S. D.; Terr. Mag. 1947 **52** 1. [16] NAISMITH, R. and BRAMLEY, E. N.; Wireless Engr. 1951 **28** 271. [17] STRANZ, D.; Ber. Dtsch. Wetterdienst U.S. Zone **14** Bad Kissingen, 1950. [18] WHALE, H. A.; J. Atmosph. Terr. Phys. 1951 **1** 233. [19] CLEGG, J. A. und DAVIDSON, I. A.; Phil. Mag. 1950 **41** 77. [20] DIEMINGER, W.; Proc. Phys. Soc. B. 1951 **64** 142. [21] APPLETON, E. V., BEYNON, W. J. G. u. PIGGOTT, W. R.; Nature, Lond. 1948 **161** 967. [22] BIBL, K. und RAWER, K.; J. Atmosph. Terr. Phys. 1951 **2** 51. [23] SCHERHAG, R.; Festvortrag Bad Kissingen 1952.

VOL. 2 1952

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 350 to 355. Pergamon Press Ltd., London

Ionospheric refraction of 81·5 Mc/s radio waves from radio stars

F. G. SMITH, M.A., Ph.D.

Cavendish Laboratory, Cambridge

(Received 4 August 1952)

ABSTRACT

Ionospheric refraction of radio waves from radio stars may cause variations in the apparent right ascensions of the stars because of horizontal gradients in the total ionisation of the ionosphere. These gradients occur during the normal diurnal changes of the *F*-region, and observations of their magnitude allow an estimate to be made of the total ionisation of the *F*-region, including the part above the maximum of ionisation density.

An account is given of measurements of the refraction of 81·5 Mc/s radio waves from four radio stars. The results are compared with calculations made from the results of ionospheric soundings, and they are found to be in agreement if certain plausible assumptions are made about the distribution of electrons in the *F*-region.

I. INTRODUCTION

Studies of the ionosphere by pulse-sounding methods are necessarily confined to the regions below maximum of ionisation of the *F*-region. Radio waves from extra-terrestrial bodies are refracted by the ionosphere so that observations of their apparent positions allow deductions to be made about the ionisation in the regions above the maximum of the *F*-region. Measurements of this type of refraction have already been made by PAYNE-SCOTT and MCCREADY [1], using solar radiation at 60 Mc/s, and by KERR, SHAIN and HIGGINS [2], using echoes from the moon at about 20 Mc/s.

The present paper describes the results of some measurements of the apparent positions of four radio stars, using a frequency of 81·5 Mc/s. The apparent regular movements of the stars are ascribed to "wedge refraction" in the ionosphere, and the results are compared with knowledge of the ionosphere obtained from ionospheric pulse-soundings.

II. WEDGE REFRACTION IN AN IONOSPHERIC LAYER

The electron density in the ionosphere is not the same over all parts of the earth at a given moment, but is a function in particular of the local solar time. The total number (n) of electrons in a vertical column of unit cross-section may therefore exhibit a gradient along an East-West line, so that the ionosphere will then behave as a thin prism, refraction in which will cause a change in the apparent right ascension of a radio star. The displacement will be a function of the local solar time, so that the refraction at the time of transit will vary throughout the year.

The value of the refraction angle θ is determined by the variation along an East-West line of the time of travel of waves through the ionosphere, and is given therefore by the gradient of $\int \mu \, dh$ where μ is the refractive index, dh is an element of height, and the integral extends through the whole ionosphere. It is well known that $1-\mu^2 = \frac{Ne^2}{\varepsilon_0 \pi m f^2}$ so that, for high frequencies where $\mu \simeq 1$, $1-\mu$ may be written:

$$1-\mu = \frac{Ne^2}{2\varepsilon_0 \pi m f^2}$$

N = electron density, m = mass of electron, f = wave frequency, ε_0 = permittivity of free space.

Under these conditions

$$\theta = \frac{d}{dx}\left\{\frac{e^2}{2\varepsilon_0 \pi m f^2}\int N\,dh\right\}$$
$$= \frac{e^2}{2\varepsilon_0 \pi m f^2}\frac{dn}{dx}.$$

Thus by observation of the apparent displacement of a radio star in right ascension it is possible to measure the horizontal gradient of n. The measurement may be made at different times throughout the day by observing the transits of several stars, and the total variation in electron content may therefore be measured.

III. Observational Method

The positional observations described in this paper were made at Cambridge with the large 3·7 m interferometer whose use in determining the positions of radio stars has already been described (Ryle, Smith and Elsmore [3]). The interferometer consisted of two large aerials spaced 105 wavelengths apart in an East-West direction. The two aerials were connected to a special type of receiver which recorded the transit of a star across the plane normal to the line of the aerials as a sinusoidally varying trace (Ryle [4]); the phase of this trace could be accurately measured to determine the apparent right ascension of the star.

VOL. 2 1952

Variable errors in the interferometer, analogous to the collimation error of an optical telescope, might cause apparent changes in the position of a radio star. These errors were minimised by observing the transits of four different stars and deducing their relative positions*. At least one of these transits occurred at a time when no refraction in right ascension was expected to occur, and this star could therefore be used as a reference to determine the apparent movements of the others.

IV. Experimental Results

Observations were made during the period July 1950 to July 1951. The results were divided into groups of about 20 separate observations, each group lasting 5–6 weeks. Of these 20 observations, a number had to be rejected because diffraction by irregularities in the ionosphere (Ryle and Hewish [5]) had reduced the accuracy, so that some groups contained only about 10 useful observations. The apparent movements caused by irregular diffraction were generally of short duration, and were distinguishable from steady refraction when observations were made over a period of about 20 min. The steady deviation of position deduced from each group is shown in Fig. 1, together with the resultant experimental error which is shown by a vertical line through each point. The full line curve in the figure will be referred to later.

The error in each of a series of observations, estimated from the knowledge of the known sources of inaccuracy, was often less than the differences between each measurement and the mean result. It is not known whether the deviations were due

* This method was adopted in a similar experiment in which an attempt was made to measure the Parallax or Proper Motion of the same four stars (Smith, 1951).

to additional small variations in the collimation error of the apparatus, or to genuine refraction effects which varied from day to day.

V. Discussion of Results

In order to decide whether the observed angular displacement could be produced by East-West gradients of the total electron content in the ionosphere it is now necessary to estimate that gradient from normal ionospheric observations.

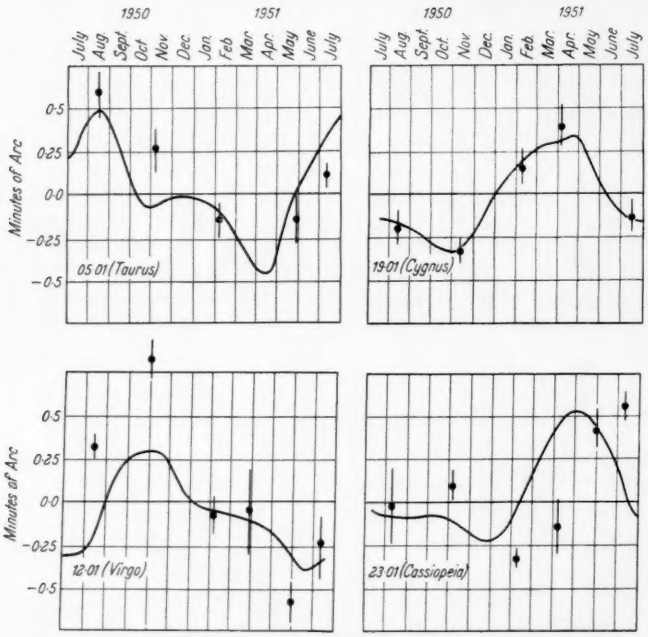


Fig. 1. Calculated (full line) and observed values of changes of apparent position caused by ionospheric wedge refraction at 81·5 Mc/s.

The total electron content (n') in a column of unit cross-section up to the maximum of the F-region could be found if the electron density $N(h)$ were known at each height h. Information about $N(h)$ is contained in the $(h' - f)$ records obtained at ionospheric observatories, but the reduction of these records to provide numerical values of $N(h)$ at all heights is tedious and is not usually performed *. Enough has been done, however, to show that the total electron content of the E-region is so small that the F-region alone need be considered in estimating n', and that the shape and thickness of the F-region often change significantly throughout the day.

Ionospheric $(h' - f)$ records are obtained regularly at Slough, and these records may be taken as characteristic of the ionosphere over South-East England where the present observations were made. Osborne [6] has used Ratcliffes approximate method [7] to analyse these records and to determine the diurnal variation of n' in the F-region on magnetically quiet days in each month of 1949, a year of sunspot maximum. In an attempt to explain the present observations, his curves were first smoothed and were then used to calculate the expected displacement of the radio stars. Although in the winter there was reasonable agreement, in the summer the calculated gradient of n' did not vary in the same way as the observed displacements. It is of interest to enquire whether there are any other observations in the light of which Osbornes results might possibly be modified.

* The Monthly Bulletins of Ionospheric Characteristics, issued by the D.S.I.R., contain values of the quantity Ym which is the semi-thickness of the parabolic distribution of ionisation approximating to the shape of the actual distribution near the maximum of ionisation. J. A. Ratcliffe has suggested to the author that, although the value of Ym is useful in the making of ionospheric forecasts, it is not useful in estimating n' since it does not properly represent the shape of the distribution at distances considerably below the maximum.

VOL. 2 1952

RATCLIFFE [8] found from an examination of results from Australia, Peru and Alaska, that the values of n' at midday were nearly always represented by a unique function (his Fig. 6) of the sun's zenith distance (x). On those few occasions when n' departed from this value there was evidence to show that there were more than the usual number of electrons above the maximum of the layer, and he made the suggestion that the total number (n) of electrons in a column of unit cross-section at midday might always vary in proportion to the value of n' in his Fig. 6. If we compare the values of n' at midday, as given by RATCLIFFE's curve for sunspot maximum, with those deduced by OSBORNE, we find that during the winter and equinoctial months OSBORNE's values vary in the same manner as RATCLIFFE's curves, although they are consistently slightly greater than his prediction. During the summer, however, OSBORNE's values are considerably lower than for the equinoctial months, whereas RATCLIFFE predicts that they should be higher (see OSBORNE's Fig. 11). It therefore seems reasonable to explore the possibility that, in both summer and winter the total number of electrons (n) over South-East England at midday varies in proportion to the value of n' shown in RATCLIFFE's Fig. 6, but that a greater proportion of them occur above the maximum in summer.

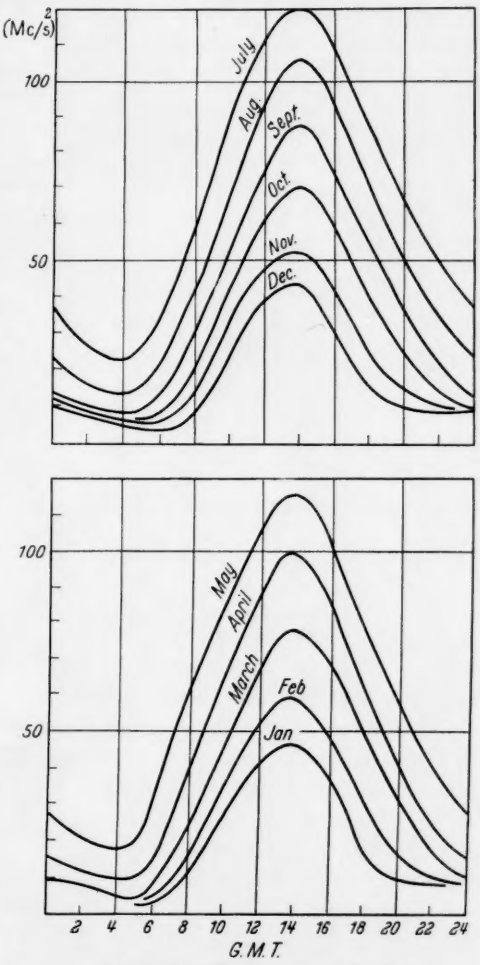


Fig. 2. Modified curves of $(fo\,F_2)^2$.

In order to make this exploration we proceed as follows. Comparison of OSBORNE's curves with the values of $fo\,F_2$ published by D.S.I.R. for 1949 shows that the values of n' throughout the day and night in winter, and near dawn and dusk in summer, bear a roughly constant ratio to the values of $(fo\,F_2)^2$ at those times. The largest discrepancies occur when the value of n' is small, and since the refraction will be least at this time these differences will be unimportant in the present investigation. Moreover the winter midday values of n' are in fair agreement with those of RATCLIFFE's Fig. 6, so we first assume that the winter day and night and the summer dawn and dusk values of n' during the period of the observations are proportional to $(fo\,F_2)^2$, the midday values being determined from RATCLIFFE's Fig. 6. The experiments described here were made between July 1950 and July 1951, so that a curve was interpolated for this purpose between RATCLIFFE's curves for sunspot maximum and minimum.

The summer midday value was taken to be that indicated by the same curve. The rest of the daytime curve in the summer months was drawn by connecting the dawn and dusk values of $(fo\,F_2)^2$ to the midday value by a smooth curve. On the hypothesis outlined here these curves would represent the diurnal variations of a quantity proportional to the number of electrons in a column of unit cross-section.

If the curves of Fig. 2 correctly represent the variation of n, then the horizontal gradient of a quantity proportional to n can be computed and compared with the observed angles of refraction of the four radio stars throughout the year. The transit of each star occurs at a different time every day; for the four stars the times may be found approximately from Fig. 3. The expected angle of refraction for the middle of each month was found from the curves of Fig. 2 using the time appropriate to each star, and these values were used to construct the curves of Fig. 1.

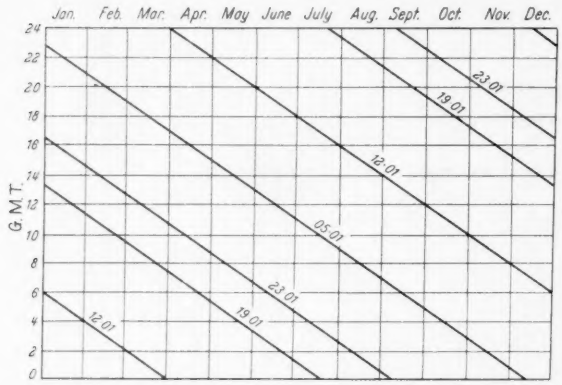


Fig. 3. The times of transit of the four radio stars.

It is evident from Fig. 1 that the experimental deviations showed the same general form as that predicted. This result is itself of considerable importance, since the form of the deviations is markedly dependent on the assumptions made in constructing the model. For example, the results for the star 05·01 show large deviations in August and in April, when the transit occurs at approximately 0800 and 1600 respectively. In Fig. 2 it is seen that these times correspond with times of rapid change of n in the model assumed: they do not however correspond to rapid changes either in the observed values of fo, or indeed in the values of n deduced by OSBORNE [6] from ionospheric pulse-sounding at Slough. The shape of the curves in Fig. 1 was deduced from the curves of Fig. 2 which show values of $(fo\,F_2)^2$. The size of the variations in the curves of Fig. 1 has been made to fit the observed results as well as possible, and from these values the absolute values of n can be deduced. It is found that, at midday, the values of n required are about four times the values of n' deduced from the results of RATCLIFFE (suitably altered to allow for an epoch halfway between sunspot maximum and minimum;) they are between two and three times the values of n' deduced for winter by OSBORNE (and adjusted for the sunspot cycle variation). At midday in summer they are about four times the values of n' deduced by OSBORNE. It is interesting to note that for a theoretical "Chapman" layer $n \simeq 3\,n'$. The difference between this result and ours is probably not significant, but it appears that in summer daytime the proportion of electrons above the maximum of the layer is greater than in winter.

VI. CONCLUSIONS

It is concluded that the variations in the observed positions of the radio stars could be explained in winter in terms of a model ionosphere in which the total electron content (n) was equal to two or three times the electron content (n') below the maximum, as deduced by OSBORNE [6] from ionospheric soundings made at Slough.

The value of n' at midday predicted by RATCLIFFE [8] is less than OSBORNEs value in winter, and in the model n at midday would be about four times this predicted value. In summer the observations would be explained if it was assumed that the radio of n/n', was increased to nearly twice the winter value, and this increase would be in agreement with suggestions made by RATCLIFFE from his study of results from other parts of the world. The winter layer would be similar to a theoretical Chapman layer; the summer one would have comparatively more electrons above the maximum.

Acknowledgements—This work was carried out at the *Cavendish Laboratory* as part of a programme of research supported by the *Department of Scientific and Industrial Research*. I wish to acknowledge the assistance given by Mr. J. A. RATCLIFFE in the interpretation of the results and in the preparation of this paper.

REFERENCES

[1] PAYNE-SCOTT, R. and MCCREADY, L. L.; Terr. Mag. 1948 **53** 429. [2] KERR, F. J., SHAIN, C. A. and HIGGINS, C. S.; Nature, Lond. 1949 **163** 310. [3] RYLE, M., SMITH, F. G. and ELSMORE, B.; M.N.R.A.S. 1950 **110** 508. [4] RYLE, M.; Proc. Roy. Soc. A 1952 **211** 351. [5] RYLE, M. and HEWISH, A.; M.N.R.A.S. 1950 **110** 381. [6] OSBORNE, B. W.; 1952 (in press). [7] RATCLIFFE, J. A.; J. Geophys. Res. 1951 **56** 463. [8] RATCLIFFE, J. A.; J. Geophys. Res. 1951 **56** 487.

VOL.
2
1952

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 356 to 360. Pergamon Press Ltd., London

Scintillation of radio stars during aurorae and magnetic storms

C. G. Little and A. Maxwell

Jodrell Bank Experimental Station, University of Manchester

(Received 23 August 1952)

ABSTRACT

During aurorae, the scintillations of radio stars are, on the average, four times more rapid than during normal conditions. This increase in the scintillation rate is believed to result from a fourfold increase in the steady drift motion of the ionospheric irregularities which cause the scintillations. More generally, it is found that the scintillation rate, and the velocity of the ionospheric irregularities, are approximately proportional to the K index of geomagnetic activity.

1. Introduction

Radio stars, when observed on metre wavelengths, are often found to "scintillate". This scintillation, or fluctuation in signal strength, is due to the diffraction of the incident radiation by irregularities in the earth's ionosphere, and is analogous to the twinkling of stars on optical wavelengths [1], [2]. It is known that the occurrence of the radio scintillations correlates with the presence of inhomogeneities in the F region, and it is now generally believed that the disturbing region is situated above the level of maximum ionisation in the F region, at a height of approximately 400 km [3], [4]. Studies of the radio scintillation phenomena therefore provide a useful method of ionospheric investigation.

This paper describes observations of radio star scintillations made on a wavelength of 3·7 m during the infrequent, but exceptionally disturbed ionospheric conditions which prevail during strong auroral and magnetic activity. It is in four main sections: (1) An analysis of the effect of auroral activity on the fluctuations of the radio stars, based on observations taken during the period 1949 May to 1951 October; (2) A more detailed correlation of radio scintillation phenomena, magnetic data, and visual and radio echo observations for the aurorae of 1951 September 25 and 1951 October 28; (3) An analysis of the correlation between the scintillation characteristics and the geomagnetic K-index; (4) Discussion of the results.

2. The Effect of Aurorae on the Scintillation of Radio Stars

Observations of the scintillation of the Cassiopeia and Cygnus radio stars during the thirty-month period from May 1949 to October 1951, have been examined for a possible correlation with auroral activity. The auroral information used in this analysis was obtained from two sources: (i) the radio echoes from aurorae [5], recorded on the continuous running meteor-survey equipment at Jodrell Bank —these records are, of course, unaffected by the presence of cloud or daylight; (ii) the summaries of visual observations of aurorae published in the Journal of the British Astronomical Association.

This analysis showed that there is no one-to-one correlation between the *amplitude* of the fluctuations and the intensity of aurorae, and that, although fluctuations were always observable during aurorae, large fluctuations were frequently observed

when no aurorae were present. The rapidity of the fluctuations, however, was found to be markedly affected by the presence of aurorae, the fluctuation rates showing, on the average, a fourfold increase during auroral activity. This effect is illustrated in Fig. 1 in which the fluctuation rates observed during eighteen aurorae are compared with those obtained in non-auroral periods.

3. Observations During the Aurorae of 1951 September 25 and 1951 October 28

Two aurorae which occurred during 1951 are of particular interest from the point of view of radio star scintillations, since it was possible to correlate changes in the scintillation characteristics with simultaneous visual, radio echo and magnetic observations.

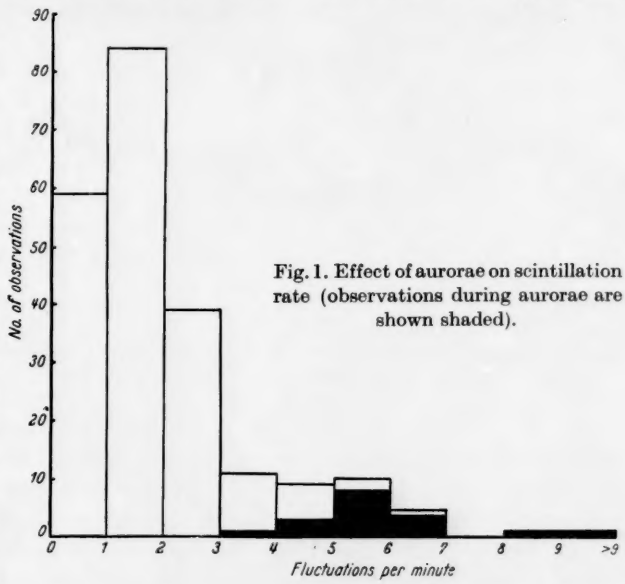


Fig. 1. Effect of aurorae on scintillation rate (observations during aurorae are shown shaded).

(a) *The aurora of 1951 September 25*—This commenced at 14.00 U.T. on 1951 September 25 and lasted for the unusually long period of 15 hrs. The accompanying storm was the most intense recorded in 1951 and persisted over the same period; radio echoes from the aurora were first observed at 15.45 U.T. and were detected at frequent intervals until 04.10 U.T. on September 26. Zenithal observations of radio star scintillations were made during the period 18.30 U.T. to 00.00 U.T. by recording the transits of the Cassiopeia and Cygnus radio stars. The following are the main points of interest arising from a comparison of the above sets of observations. (i) Stable auroral arcs do not appear to influence radio scintillation phenomena. This was shown firstly, when a stable auroral arc moved slowly across the Cygnus region without affecting the radio scintillations, and secondly, when the zenithal fluctuations remained unchanged by the appearance of an auroral arc low on the northern horizon. (ii) Rapidly varying auroral formations have a very marked effect upon radio scintillation characteristics. This was shown most clearly at 20.30 U.T. when a stable auroral arc suddenly broke up into a series of pulsating rays and diffuse patches, the flaming aurora then observed being the brightest recorded in Great Britain for 20 years. Simultaneously the radio star scintillations increased in amplitude from approximately 10% modulation of the mean source intensity, to intense fluctuations with peaks up to 300% of the mean source intensity; the fluctuation rate also increased from 1 per minute to approximately 8 per minute at the same time. With the onset of the flaming aurora the magnetic record showed

VOL. 2 1952

a large pulse (after an hour of comparative quiescence during the stable auroral formation) and the radio echoes changed in type from diffuse to intense discrete echoes.

(b) *The aurora of 1951 October 28*—This was of unusual interest in that a marked decrease in the intensity of the extra-terrestrial radio emission was observed, coincident in time with the maximum activity of the aurora. On this occasion, observations weremade with three equipments operating on 81·5 Mc/s and situated some 3 km apart. All the records show a sudden decrease of the order of 10% in the extra-terrestrial signal, followed by a slow rise to the original value. The effect is illustrated in Fig. 2. which reproduced the seconds for two of the equipments.

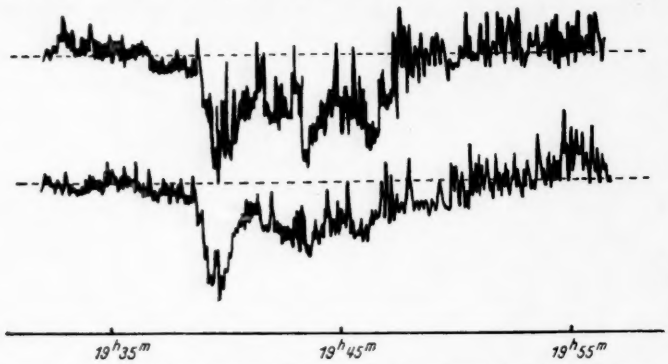


Fig. 2. Records taken by receivers spaced 3 km apart showing decrease in extra-terrestrial radio signal during aurora of 1951 October 28.

There are two possible mechanisms which would serve to explain the decrease in the intensity of the extra-terrestrial radio waves. It is possible that the auroral ionisation was sufficiently intense to reflect some of the incident radiation away from the observer; alternatively the incoming radiation may have been partially absorbed by an increase in the ionisation at heights of the order of 70 km. The former explanation is thought to be the more probable since (i) 72 Mc/s radar echoes from auroral formations were observed at this time, indicating the presence of regions of sufficiently high electron density to reflect some of the incident radiation; (ii) visual observations indicate that the auroral formations at the time of the decrease were at a height of several hundred kilometres; at such heights the absorption of 80 Mc/s radiation by collision damping would be negligible.

Some information as to the lateral dimensions of the disturbing region may be derived from the spaced receiver observations. Thus, the fact that receivers spaced some 3 km apart recorded a similar decrease indicates that the dimensions of the region were greater than 3 km. Also, on the assumption that the duration of the decrease was associated with the motion of some quasi-stable disturbance across the line of sight, it is possible to make an estimate of the lateral dimensions of the region from a knowledge of its velocity and the time taken by it to pass the observer. During the aurora, the average wind velocity in the upper ionosphere (estimated from the fluctuation rate [6]) was of the order of 400 m/sec; since the period of the decrease was approximately 5 min, this suggests that the disturbing region had lateral dimensions of the order of 120 km. If, however, the duration of the decrease were associated with the growth and decay of an extended disturbing region, it is not possible to deduce the lateral dimensions of the disturbing region in this way.

4. The Correlation between the Scintillation Rate and the Geomagnetic K-Index

It has been shown above that, although the amplitude of the fluctuations is relatively unaffected by the presence of aurorae, the radio scintillation rate is considerably faster during an aurora than during normal conditions. Aurorae are always accompanied by severe geomagnetic disturbances, and it is therefore of interest to compare the fluctuation rates with the K-index of geomagnetic activity, during both disturbed and non-disturbed periods. A detailed analysis of this type has been carried out, and the results are shown in Fig. 3. From this it will be seen that there is significant correlation between the two phenomena. Statistical analysis of the results shows that the correlation coefficient is 0·74 and that the probability that the two phenomena are unrelated is less than one in 1000.

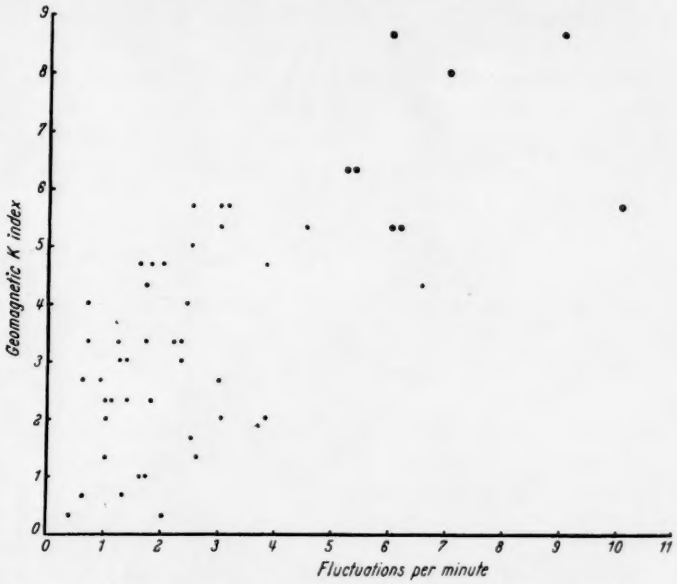


Fig. 3. Scintillation Rate v. Geomagnetic K-index. (The K-indices used here are the average of the corresponding Abinger, Eskdalemuir and Lerwick 3-hour figures.) O Observations during aurorae. ● Observations during non-auroral periods.

5. Discussion

It is known that radio scintillations are due to the diffraction of the incident radiation by irregularities in the F region. These irregularities produce a non-uniform distribution of radio intensity across the ground; any motion of this diffraction pattern, or changes in its structure with time give rise to fluctuations in the radio intensity recorded at a fixed point on the ground. An increase in the speed of fading, such as is observed during aurorae, could therefore be due to (1) a decrease in the average dimensions of the irregularities in the diffraction pattern, (2) an increase in the velocity of the pattern across the ground, and (3) an increase in the rate at which the diffraction pattern changes with time.

It is possible to differentiate between these three mechanisms by using spaced receivers [3], [6]. This technique enables the velocity, dimensions and rate of change of the diffraction pattern to be determined from a comparison of the records. Such observations show that the rate of change of the diffraction pattern is too slow to contribute appreciably towards the fluctuation rate, and that over a 20:1 range, the fluctuation rate is proportional to the velocity of the diffraction pattern across the ground. This indicates, therefore, that the changes in fluctuation rate are due to changes in the velocity of the diffraction pattern, and not to changes

VOL. 2 1952

in its dimensions or stability. During aurorae the time differences between records taken at the three triangularly situated receiving equipments were too small to enable the pattern velocity to be measured accurately, but the fact that the correlation between the fluctuation records was still maintained indicates that the scale of the irregularities remained appreciably the same.

We may conclude therefore, that the high fluctuation rates observed during magnetically disturbed periods and during aurorae correspond to an increase in the velocity of the ionospheric winds, and that, in general, the velocity of these winds is proportional to the magnetic K-index.

Acknowledgements—This work forms part of a research programme made possible by financial assistance from the *Department of Scientific and Industrial Research*. The authors wish to thank Professor A. C. B. Lovell, director of the Jodrell Bank Experimental Station, for his helpful advice and criticism; they are also indebted to their colleague Dr. G. S. Hawkins for permission to refer to unpublished data on radio echoes from the aurorae, and to Mr. J. Paton of Edinburgh University for information about the visual aspects of the aurorae of 1951 September 25 and October 28.

References

[1] Little, C. G.; M.N.R.A.S. 1951 **111** 289. [2] Hewish, A.; Proc. Roy. Soc. A. 1951 **209** 81. [3] Little, C. G. and Maxwell, A.; Phil. Mag. 1951 **42** 267. [4] Ryle, M. and Hewish, A.; M.N.R.A.S. 1950 **110** 381. [5] Aspinall, A. and Hawkins, G. S.; J. Brit. Astr. Assoc. 1950 **60** 130. [6] Maxwell, A. and Little, C. G.; Nature, Lond. 1952 **169** 746.

RESEARCH NOTES

The Quasi-Transverse (Q.T.) approximation to Appleton's magneto-ionic equation

ABSTRACT

It is shown that the usually accepted form of the quasi-transverse (Q.T.) approximation to APPLETON's magneto-ionic equation is incorrect and the correct form is given.

When an electro-magnetic wave passes through a medium containing free electrons, the refractive index (μ) and the absorption coefficient (χ) are given by APPLETON's well-known expression

$$(\mu - i\chi)^2 = 1 - \frac{x}{1 - iz - \dfrac{y_T^2}{2(1-x-iz)} \pm \sqrt{\dfrac{y_T^4}{4(1-x-iz)^2} + y_L^2}} \tag{1}$$

(omitting the Lorentz term).

The meaning of the symbols in the equation are explained in the original paper by APPLETON [1]. Eq. (1) is complicated and leads to cumbersome expressions, but BOOKER [2] has shown how it is often convenient to use, instead, one of two simpler approximate forms. Much use has been made of these approximations in the literature. The Q.T. approximation is valid when

$$\frac{y_T^4}{4y_L^2} \gg (1-x)^2 + z^2$$

and when it holds, the approximation to (1), for the positive sign in the denominator (the ordinary wave), is usually taken to be

$$(\mu - i\chi)^2 = 1 - \frac{x}{1 - iz}. \tag{2}$$

The following two cases are then of especial interest.

(A) when $z \ll 1$ and $\chi^2 \ll \mu^2$

$$\mu = \sqrt{1-x} \tag{3}$$

and

$$\chi = \frac{xz}{2\sqrt{1-x}}. \tag{4}$$

(B) when $z \ll 1$ and $x = 1$

$$(\mu - i\chi)^2 = -iz$$

or

$$\mu = \chi = \sqrt{\frac{z}{2}}. \tag{5}$$

The purpose of this note is to point out that the usually accepted form (2) of the Q.T. approximation is incorrect. A first order term has been omitted in its derivation, as follows.

The denominator in the last term of eq. (1) is

$$1 - iz + \frac{y_T^2}{2(1-x-iz)}\left\{\sqrt{1 + \frac{4y_L^2(1-x-iz)^2}{y_T^4}} - 1\right\}$$

and when

$$\frac{y_T^4}{4y_L^2} \gg (1-x)^2 + z^2$$

it becomes

$$\begin{aligned} &1 - iz + \frac{y_T^2}{2(1-x-iz)}\left\{1 + \frac{2y_L^2(1-x-iz)^2}{y_T^4} - 1\right\} \\ &\quad = 1 - iz + (1-x-iz)\frac{y_L^2}{y_T^2} \\ &\quad = 1 - iz + (1-x-iz)\cot^2\theta \end{aligned} \tag{6}$$

VOL. 2 1952

where θ is the angle between the wave-normal and the direction of the imposed magnetic field. The correct form of the Q.T. approximation is therefore

$$(\mu - i\chi)^2 = 1 - \frac{x}{1 - iz + (1 - x - iz)\cot^2\theta} \tag{7}$$

and this expression should replace eq. (2). The two important special cases can now be re-examined

(A) when $z \ll 1$ and $\chi^2 \ll \mu^2$

$$\mu = \sqrt{\frac{(1-x)(1+\cot^2\theta)}{1+(1-x)\cot^2\theta}} = \operatorname{cosec}\theta\sqrt{1-x} \quad \text{when} \quad (1-x) \ll 1 \tag{8}$$

and

$$\chi = \frac{xz}{2\mu\{1+(1-x)\cot^2\theta\}} = \frac{xz\operatorname{cosec}\theta}{2\sqrt{1-x}} \quad \text{when} \quad (1-x) \ll 1 \tag{9}$$

(B) when $z \ll 1$ and $x = 1$

$$(\mu - i\chi)^2 = -iz\operatorname{cosec}^2\theta$$

or

$$\mu = \chi = \operatorname{cosec}\theta\sqrt{\frac{z}{2}}. \tag{10}$$

The eqs. (8), (9) and (10) should replace eqs. (3), (4) and (5). The additional factor cosec θ can, under certain circumstances, make an important difference to the calculated amount of deviative absorption which occurs near the level of reflection of a wave incident vertically on the ionosphere. A detailed examination of the consequences is being made.

References

[1] Appleton, E. V.; J. Inst. Elect. Engrs. 1932 **71** 430, 642–650. [2] Booker, H. G.; Proc. Roy. Soc. A (No. 870) 1935 **150** 267–285.

Cavendish Laboratory, Cambridge, 13 August 1952 J. D. Whitehead

A physical interpretation of the corrected quasi-transverse approximation for the ordinary wave

The corrected Q.T.-approximation as given above by Whitehead can be given an approximate physical interpretation. It is well known [1] that the ordinary wave is deviated in the magnetic meridian plane, and in the case of no collisions, is reflected when the Poynting flux is perpendicular to the earth's magnetic field and where both the electric vector and the polarization vector vibrate along the field and therefore are unaffected by it. The introduction of a small amount of damping does not seriously alter this picture.

The formulation of the Appleton-Hartree equations is such that the total attenuation phase- and group-path, are found by integrating vertically, *i.e.* perpendicularly to the stratifications. As shown by Whitehead, the absorption coefficient in the Q.T.-region is cosec θ times the no-field value. This seems explicable by the increased group path ds near the reflection point, since from simple geometry we have $ds = \operatorname{cosec}\theta\, dh$. Likewise the group delay near the reflection point is increased over the no field value by the same factor. This is also apparent from Fig. 2 in a recent paper by Shinn and Whale [2].

We can therefore say approximately that in the Q.T.-region the ordinary wave is propagated unaffected by the field, but that we must evaluate the integrals along the group path which is strongly influenced by the field.

References

[1] Booker, H. G.; An Outline of the Magneto-Ionic Theory, Cornell University. [2] Shinn, D. H. and Whale, H. A.; J. Atmosph. Terr. Phys. 1952 **2** 85.

Cavendish Laboratory, Cambridge, 13 August 1952 F. Lied

Journal of Atmospheric and Terrestrial Physics, 1952, Vol. 2, pp. 363 to 364. Pergamon Press Ltd., London

Indices of geomagnetic activity

of the Observatories ABINGER (*Ab*), ESKDALEMUIR (*Es*) and LERWICK (*Le*)

July and August 1952

The figures given on pages 363 to 364 represent the *K*-indices for three-hour intervals, beginning with 00–03 hrs for the first and ending with 21–24 hrs for the eighth figure.

July 1952

Day	*Ab* Range for $K = 9$: 500 γ		Sum	*Es* Range for $K = 9$: 750 γ		Sum	*Le* Range for $K = 9$: 1000 γ		Sum	Day
	K-Indices			*K-Indices*			*K-Indices*			
1	1122	2255	20	1122	2145	18	1112	1145	16	1
2	3223	2322	19	2223	2222	17	2113	2211	13	2
3	2223	3442	22	2212	3432	19	2212	3332	18	3
4	2223	3412	19	1223	3312	17	1213	3312	16	4
5	3444	5554	34	3434	4553	31	2443	5665	35	5
6	3453	3322	25	3453	3332	26	3453	3222	24	6
7	3232	2320	17	3222	2320	16	2222	2321	16	7
8	1132	4333	20	1122	3323	17	1121	3323	16	8
9	3334	3443	27	3334	3432	25	3233	3433	24	9
10	2333	3533	25	2323	3543	25	2232	3543	24	10
11	3333	3330	21	3332	3330	20	4331	3220	18	11
12	2223	3222	18	2222	2212	15	3222	2112	15	12
13	1223	3432	20	1112	3431	16	1212	3321	15	13
14	3322	3423	22	2322	3423	21	3322	3323	21	14
15	3423	4433	26	2423	3333	23	2423	3333	23	15
16	3321	3331	19	2222	3321	17	2221	2221	14	16
17	0222	2333	17	0222	2333	17	1222	2222	15	17
18	1323	3333	21	1212	3222	15	1322	2212	15	18
19	1122	3232	16	1112	3232	15	1112	2122	12	19
20	3225	5645	32	2225	5534	28	2224	4535	27	20
21	3444	5444	32	3444	5444	32	4543	5534	33	21
22	4333	3334	26	3323	3333	23	4323	2233	22	22
23	2332	3334	23	2222	2323	18	3221	2323	18	23
24	3223	3333	22	3202	3333	19	3112	3223	17	24
25	1112	3433	18	1122	3432	18	2122	3333	19	25
26	3333	2311	19	3332	2211	17	4422	2211	18	26
27	2122	3324	19	2121	2224	16	2111	2224	15	27
28	2222	2231	16	2111	1220	10	2211	1221	12	28
29	0111	2221	10	0001	2121	7	0111	2221	10	29
30	1012	2111	9	2110	0111	7	1110	1111	7	30
31	2244	3442	25	2133	3432	21	2133	3331	19	31

August 1952

Day	*Ab* Range for $K = 9$: 500 γ		Sum	*Es* Range for $K = 9$: 750 γ		Sum	*Le* Range for $K = 9$: 1000 γ		Sum	Day
	K-Indices			*K-Indices*			*K-Indices*			
1	2322	2332	19	2221	2232	16	2321	2232	17	1
2	2211	3244	19	2111	3244	18	3111	2233	16	2
3	2333	5532	26	2332	5432	24	2433	5533	28	3
4	3332	2444	25	3332	3344	25	3332	2334	23	4
5	4223	2332	21	3222	2332	19	4222	1332	19	5
6	4433	3432	26	4423	2322	22	5432	2323	24	6
7	4343	3233	25	4231	2223	19	5221	2223	19	7
8	3222	3332	20	3222	3232	19	3222	2122	16	8
9	0222	2233	16	1112	1133	13	1112	2233	15	9
10	4222	4443	25	4222	4432	23	4322	4433	25	10
11	4113	3334	22	4212	4224	21	4312	3224	21	11
12	5233	3443	27	3233	3442	24	6322	3332	24	12
13	1121	1311	11	1122	2211	12	1221	2211	12	13
14	2222	1221	14	1211	2210	10	1221	2121	12	14
15	2112	2233	16	2101	1233	13	2111	2232	14	15
16	2322	2211	15	2221	1111	11	2221	1111	11	16
17	3344	5454	32	2344	4454	30	2234	4343	25	17
18	5333	4555	33	5322	3444	27	5432	3454	30	18
19	4324	3431	24	3323	3321	20	3323	2322	20	19
20	3433	3334	26	3333	3323	23	3333	2224	22	20
21	2124	2321	17	2012	2310	11	1112	2321	13	21
22	1123	2223	16	0112	2222	12	1111	1223	12	22
23	4223	3332	22	4212	2321	17	4312	2221	17	23
24	1132	3321	16	1122	2320	13	1121	1220	10	24
25	0123	3211	13	0012	3211	10	0112	3212	12	25
26	1213	2222	15	1112	2222	13	1112	2222	13	26
27	3333	2221	19	3232	1121	15	3222	1221	15	27
28	2211	1211	11	2210	1211	10	2211	1111	10	28
29	2122	3444	22	2122	2444	21	2112	3455	23	29
30	4442	3333	26	4542	2332	25	6642	2333	29	30
31	1222	2223	16	1212	2123	14	1212	2125	16	31

VOL.
2
1952